GenCore version 5.1.6
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: February 5, 2003, 12:42:26 ; Search time 7.20872 Seconds
(without alignments)
200.219 Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MPPIIELSP[***]MPPP.....IIIEEYIPEL_EYPIL_NC 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262774 seqs, 206,2xx residues

Total number of hits satisfying chosen parameters: 262774

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /.../.../5_..._1.pep:*
2: /.../.../6_..._1.pep:*
3: /.../.../7_..._1.pep:*
4: /.../.../8_..._1.pep:*
5: /.../.../9_..._1.pep:*
6: /.../.../10_..._1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	260	100.0	191	4	US-xx-xxx-x	Sequence 1, Appli
2	259	99.6	191	4	US-xx-xxx-x	Sequence 5, Appli
3	257	98.8	191	4	US-xx-xxx-x	Sequence 4, Appli
4	257	98.8	191	4	US-xx-xxx-x	Sequence 4, Appli
5	255	98.1	217	4	US-xx-xxx-x	Sequence 10, App
6	255	98.1	217	4	US-xx-xxx-x	Sequence 1, App
7	255	98.1	217	4	US-xx-xxx-x	Sequence 10, App
8	255	98.1	217	4	US-xx-xxx-x	Sequence 11, App
9	255	98.1	217	4	US-xx-xxx-x	Sequence 1, App
10	255	98.1	241	4	US-xx-xxx-x	Sequence 25, App
11	255	98.1	274	4	US-xx-xxx-x	Sequence 71, App
12	255	98.1	300	4	US-xx-xxx-x	Sequence 1, App
13	249	95.8	191	4	US-xx-xxx-x	Sequence 1, App
14	248	95.4	217	1	US-xx-xxx-x	Sequence 4, App
15	247	95.0	217	1	US-xx-xxx-x	Sequence 51, App
16	247	95.0	217	2	US-xx-xxx-x	Sequence 51, App
17	247	95.0	217	2	US-xx-xxx-x	Sequence 4, App
18	246	94.6	191	4	US-xx-xxx-x	Sequence 16, App
19	246	94.6	191	4	US-xx-xxx-x	Sequence 16, App
20	244	93.8	191	4	US-xx-xxx-x	Sequence 2, App
21	243	93.5	217	2	US-xx-xxx-x	Sequence 24, App
22	243	93.5	217	2	US-xx-xxx-x	Sequence 24, App
23	243	93.5	217	2	US-xx-xxx-x	Sequence 24, App
24	242	93.1	191	1	US-xx-xxx-x	Sequence 8, Appli
25	164	63.1	179	3	US-xx-xxx-x	Sequence 1, App
26	164	63.1	179	4	US-xx-xxx-x	Sequence 1, App
27	163.5	62.9	191	1	US-07-xxx-x	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-xx-xxx-x
; Sequence 1, Application US/xx/xxxxxx
; Patent No. 5xxxxxx
; GENERAL INFORMATION:
; APPLICANT: [redacted]
; APPLICANT: [redacted]
; APPLICANT: [redacted]
; TITLE OF INVENTION: DNA Expression Systems
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [redacted]
; STREET: [redacted]
; CITY: [redacted]
; STATE: [redacted]
; COUNTRY: USA
; ZIP: [redacted]
; COMPUTER READABLE FORM:
; MEDIUM TYPE: [redacted]
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: [redacted]
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/xx/xxxxxx
; FILING DATE: 24-[redacted]
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [redacted]
; FILING DATE: [redacted]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [redacted]
; FILING DATE: [redacted]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [redacted]
; FILING DATE: [redacted]
; ATTORNEY/AGENT INFORMATION:
; NAME: [redacted]
; REGISTRATION NUMBER: [redacted]
; REFERENCE/DOCKET NUMBER: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [redacted]
; TELEFAX: 415/[redacted]
; TELEX: [redacted]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [redacted] base pairs
; TYPE: [redacted]

```
        1 1.0%; Gaps    0;
US-08-284-878-5

  Query Match             100.0%;  Score 260;  DB 1;  Length 192;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-39;
  Matches  48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 48
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 48


RESULT 2
US-09-284-878-5
; Sequence 5, Application US/09284878
; Patent No. 6541575
; GENERAL INFORMATION:
;  APPLICANT:  Saccardo, Martha Gabriela
;  APPLICANT:  Saldana, Hugo Barrera
;  APPLICANT:  Salvado, Jose Maria Viader
;  TITLE OF INVENTION:  Genetically Modified Methylotrophic P. pastoris Yeast for the
;  TITLE OF INVENTION:  Production and Secretion of the Human Growth Hormone
;  FILE REFERENCE:  1829.001Dus00
;  CURRENT APPLICATION NUMBER:  US/09/284,878
;  CURRENT FILING DATE:  1999-07-21
;  PRIOR APPLICATION NUMBER:  PCT/MX97/00008
;  PRIOR FILING DATE:  1997-10-24
;  NUMBER OF SEQ ID NOS:  9
;  SOFTWARE:  Patentin Ver. 2.1
; SEQ ID NO 5
;  LENGTH: 192
;  TYPE: PRT
;  ORGANISM: Homo sapiens
US-09-284-878-5

  Query Match             99.1%;  Score 259;  DB 1;  Length 192;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-39;
  Matches  48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 48
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 48


RESULT 3
US-08-383-621-4
; Sequence 4, Application US/08383621
; Patent No. 5951972
; GENERAL INFORMATION:
;     APPLICANT:  Daley, Michael J.
;     APPLICANT:  Buckwalter, Brian L.
;     APPLICANT:  Cady, Susan M.
;     APPLICANT:  Shieh, Hong-Ming
;     APPLICANT:  Bohlen, Peter
;     APPLICANT:  Sodum, Andrew P.
;     TITLE OF INVENTION:  Stabilization of Somatotropins And Other
;     TITLE OF INVENTION:  Proteins By Modification of Cysteine Residues
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dr. Patricia J. Hewitt
;       STREET:  1937 West Main Street, P.O. Box 30
;       CITY:  Stamford
;       STATE:  Connecticut
;       COUNTRY:  U.S.A.
;       ZIP:  06904-0026
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  08/383,621
;       FILING DATE:  06-FEB-1995
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/[illegible]
;       FILING DATE:  25-SEP-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Isovdos, Estelle J.
;       REGISTRATION NUMBER:  31,113
;       REFERENCE/DOCKET NUMBER:  [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  203-321-2556
;       TELEFAX:  203 321 2971
;       TELEX:  203 321-3401-4679
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  194 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-383-621-4

  Query Match             98.1%;  Score 257;  DB 1;  Length 194;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-39;
  Matches  48;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 EPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 49
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 EPITPCSRLEFNAMLRAHRLHLAFDTYQEFDEAYIPKEQKYSFLQNPQ 51


RESULT 4
US-08-459-906-4
; Sequence 4, Application US/08459906
; Patent No. 6010999
;  GENERAL INFORMATION:
;     APPLICANT:  Daley, Michael J.
;     APPLICANT:  Buckwalter, Brian L.
;     APPLICANT:  Cady, Susan M.
;     APPLICANT:  Shieh, Hong-Ming
;     APPLICANT:  Bohlen, Peter
;     APPLICANT:  Sodum, Andrew P.
;     TITLE OF INVENTION:  Stabilization of Somatotropins and Other
;     TITLE OF INVENTION:  Proteins by Modification of Cysteine Residues
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  American Cyanamid Company
;       STREET:  One Cyanamid Plaza
;       CITY:  Wayne
;       STATE:  New Jersey
;       COUNTRY:  U.S.A.
;       ZIP:  07470-8426
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Patentin Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/[illegible]
;       FILING DATE:  02-JUN-1995
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Webster, Barry C.
;       REGISTRATION NUMBER:  34,758
;       REFERENCE/DOCKET NUMBER:  [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  201 831 4247
;       TELEFAX:  201 831 [illegible]
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  194 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-459-906-4
```

; STATE: Texas
; COUNTRY: USA
; ZIP: 77211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/785,771
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/589,028
; FILING DATE: 19-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn, Steven L.
; REGISTRATION NUMBER: 47,643
; REFERENCE/DOCKET NUMBER: 0150.51
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 10
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-785-771-10

Query Match 98.1%; Score 255; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP
Db 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 51

RESULT 8
US-08-785-771-10
; Sequence 10, Application US/09/785,771
; Patent No. 6422449
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W.
; APPLICANT: Beck, Fernando J.
; APPLICANT: Bucio, J. Fernando
; APPLICANT: Kosteiin, Robert A.
; TITLE OF INVENTION: METHANOTROPHIC P. PASTORIS STRAINS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/785,628
; FILING DATE: 05-FEB-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/180,574
; FILING DATE: 07-FEB-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Chinn, Edwin P.
; REGISTRATION NUMBER: 41,545
; REFERENCE/DOCKET NUMBER: 0150.55

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415.904.6500
; TELEFAX: 415.904.6520
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..217
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 94..115
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..217
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 142..217
; OTHER INFORMATION: /note= "The peptides above are
; OTHER INFORMATION: separated by a linker."
US-08-785-628-11

Query Match 98.1%; Score 255; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP
Db 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 51

RESULT 9
US-09-423-100-1
; Sequence 1, Application US/09/423100
; Patent No. 6482927
; GENERAL INFORMATION:
; APPLICANT: Sivasubramanian, Murthy Sundaram
; APPLICANT: Suldana, Hugo Barrera
; APPLICANT: Salcedo, Jose Maria Vicente
; TITLE OF INVENTION: Genetically Modified Methylotrophic Pichia pastoris Yeast Strain
; TITLE OF INVENTION: Production and Secretion of Human Serum Albumin
; FILE REFERENCE: IRI-0100PCT
; CURRENT APPLICATION NUMBER: US/09/423,100
; CURRENT FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: PCT/US98/09054
; PRIOR FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-1

Query Match 98.1%; Score 255; DB 1; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.2e-20;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP
Db 27 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 76

RESULT 10
US-09-423-100-25
; Sequence 25, Application US/09/423,100
; Patent No. 6-015-5
; GENERAL INFORMATION:
; APPLICANT: SYNCHELLE PIERRE

; CAN; Kitayama
; MAN, Jae-Wook
; PAR, Charles
; YAN, Dae-Sok
; LEE, Joo-Woo
; SEONG, Baik-Lin
; TITLE OF INVENTION: Process for preparing recombinant proteins using highly
; efficient expression vector I in Saccharomyces cerevisiae
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY, P.C.
; STREET: Suite 1201, 900 Chapel Street
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06510-2807
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 1.44 Mb storage
; COMPUTER: IBM
; OPERATING SYSTEM: WINDOWS 95/NT
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/124,820B
; FILING DATE: 21-May-1999
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-423-100-24

Query Match 98.1%; Score 255; DB 4; Length 241;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFETDESKEINAMEPARPCHLIAPLYLFFEAYIEKLKYSFLNE 48
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EFETDESKEINAMEPARPCHLIAPLYLFFEAYIEKLKYSFLNE 48

RESULT 2
US-09-423-100-25
; Sequence 25, Application US/09/423,100
; Patent No. 6511807
; GENERAL INFORMATION:
; APPLICANT: Nowakowski, Christopher B.
; APPLICANT: Bolton, Philippe A.
; APPLICANT: Xu, Yiliu(Williamson, Karl R.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Illingen, Adrian E.
; APPLICANT: Quando, Christian
; APPLICANT: Kroiss, Fred
; APPLICANT: McGarry, Dennis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION SYSTEM FOR GLYCOSYLATED
; TITLE OF INVENTION: SECRETED POLYPEPTIDES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: September-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,260
; FILING DATE: 06-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hildebrand, Kevin D.
; REGISTRATION NUMBER: 36,197
; REFERENCE/DOCKET NUMBER: 1758.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-423-100-25

Query Match 98.1%; Score 255; DB 4; Length 234;
Best Local Similarity 100.0%; Pred. No. 2.5e-24;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EFETDESKEINAMEPARPCHLIAPLYLFFEAYIEKLKYSFLNE 48
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 27 EFETDESKEINAMEPARPCHLIAPLYLFFEAYIEKLKYSFLNE 74

RESULT 2
US-09-423-100-25
; Sequence 25, Application US/09/423,100
; Patent No. 6511807
; GENERAL INFORMATION:
; APPLICANT: Nowakowski, Christopher B.
; APPLICANT: Bolton, Philippe A.
; APPLICANT: Xu, Yiliu(Williamson, Karl R.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Illingen, Adrian E.
; APPLICANT: Quando, Christian
; APPLICANT: Kroiss, Fred
; APPLICANT: McGarry, Dennis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION SYSTEM FOR GLYCOSYLATED
; TITLE OF INVENTION: SECRETED POLYPEPTIDES
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: September-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,744
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,260
; FILING DATE: 06-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hildebrand, Kevin D.
; REGISTRATION NUMBER: 36,197
; REFERENCE/DOCKET NUMBER: 1758.1

; COMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-784-582-3

Query Match 98.9%; Score 265; DB 3; Length 96;
Best Local Similarity 100.0%; Pred. No. 4.1e-24;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 51
||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50

RESULT 2
US-08-465-461-3
; Sequence 3, Application US/08465461
; Patent No. 5684444
; GENERAL INFORMATION:
; APPLICANT: CHAPPEL, Scott
; TITLE OF INVENTION: Human Growth Hormone to stimulate hematopoiesis and immune recons
; TITLE OF INVENTION: after hematopoietic stem cell transplantation in humans
; FILE REFERENCE: CHAPPELL 6.1
; CURRENT APPLICATION NUMBER: US/08/465,461
; CURRENT FILING DATE: 1995-12-27
; PRIOR APPLICATION NUMBER: 60/112,068
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Version 3.1
; SEQ ID NO 3
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-465-461-3

Query Match 95.8%; Score 249; DB 4; Length 191;
Best Local Similarity 97.9%; Pred. No. 1.4e-20;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 51
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50

RESULT 3
US-08-287-756-4
; Sequence 4, Application US/08287756
; Patent No. 5714352
; GENERAL INFORMATION:
; APPLICANT: OLSEN, Eric A.
; TITLE OF INVENTION: Human Growth Hormone
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,756
; FILING DATE: 09-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE: PROTEIN
US-08-287-756-4

Query Match 95.8%; Score 249; DB 4; Length 217;
Best Local Similarity 97.9%; Pred. No. 1.6e-20;
Matches 47; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 51
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 76

RESULT 4
US-08-469-486-51
; Sequence 51, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thogersen, Hans Christian
; APPLICANT: Bolten, Ilse Lau
; APPLICANT: Luzzago, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul L. Clark
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid

Migration protein search, using sw model

run on: February 5, 2003, 12:42:21 ; Search time 20.25 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MET[illegible]

Scoring table: BLOSUM62
Gapop [illegible] , Gapext 1.5

Searched: 9847 seqs, 1325[illegible] residues

Total number of hits satisfying chosen parameters: 9847

Minimum DB seq length:
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]
7: [illegible]
8: [illegible]
9: [illegible]
10: [illegible]
11: [illegible]
12: [illegible]
13: [illegible]
14: [illegible]
15: [illegible]
16: [illegible]
17: [illegible]
18: [illegible]
19: [illegible]
20: [illegible]
21: [illegible]
22: [illegible]
23: [illegible]
24: [illegible]
25: [illegible]
26: [illegible]
27: [illegible]
28: [illegible]
29: [illegible]
30: [illegible]
31: [illegible]
32: [illegible]
33: [illegible]
34: [illegible]
35: [illegible]
36: [illegible]
37: [illegible]
38: [illegible]
39: [illegible]
40: [illegible]
41: [illegible]
42: [illegible]
43: [illegible]
44: [illegible]
45: [illegible]

Prot. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

result No.	Query match	Length	DB ID	Description
1	100	49	AAY12[illegible]	Human growth hormo
2	100	49	AAB12[illegible]	Human growth hormo
3	100	191	AAY128[illegible]	5-HT[illegible]-protein ri
4	100	191	AAW222[illegible]	Human anti-androge
5	100	191	AAB[illegible]	Human growth hormo
6	100	191	AAY128[illegible]	'Dimeric' protein,
7	100	192	AAB[illegible]	Human growth hormo
8	100	192	AAW[illegible]	Human anti-androg
9	100	200	AAB[illegible]	Human beta nerve g
10	100	262	AAB[illegible]	Human beta nerve g
11	100	[illegible]	AAB[illegible]	Human growth hormo
12	100	[illegible]	AAB[illegible]	Human growth hormo
13	100	[illegible]	AAB[illegible]	Human growth hormo
14	100	[illegible]	AAB[illegible]	Human growth hormo
15	100	[illegible]	AAY[illegible]	Human growth hormo
16	100	[illegible]	AAW[illegible]	Human growth hormo
17	100	[illegible]	AAB[illegible]	Human growth hormo
18	100	[illegible]	AA[illegible]	Human growth hormo
19	100	[illegible]	AAW[illegible]	Human growth hormo
20	100	[illegible]	AAY[illegible]	Human growth hormo
21	100	[illegible]	AAY[illegible]	Human growth hormo
22	100	[illegible]	AAY[illegible]	Human growth hormo
23	100	[illegible]	AAY[illegible]	Human growth hormo
24	100	[illegible]	AA[illegible]	Human growth hormo
25	100	[illegible]	AA[illegible]	Human growth hormo
26	100	[illegible]	AAB[illegible]	Met Asp human growth h
27	100	[illegible]	AAY[illegible]	Human growth hormo
28	100	[illegible]	AAB[illegible]	Sequence 1, Appl
29	100	[illegible]	AAB[illegible]	Sequence 1, Appl
30	100	[illegible]	AAB[illegible]	Sequence 1, Appl
31	100	[illegible]	AAB[illegible]	Human growth hormo
32	100	[illegible]	AAW[illegible]	Sequence 1, Appl
33	100	[illegible]	AAB[illegible]	Human growth hormo
34	100	[illegible]	AAW[illegible]	Human growth hormo
35	100	[illegible]	AAY[illegible]	Human growth hormo
36	100	[illegible]	AAB[illegible]	Human growth hormo
37	100	[illegible]	AAB[illegible]	Human growth hormo
38	100	[illegible]	AAW[illegible]	Human growth hormo
39	100	[illegible]	AAB[illegible]	Sequence 1, Appl
40	100	[illegible]	AA[illegible]	Human growth hormo
41	100	[illegible]	AA[illegible]	Human growth hormo
42	100	[illegible]	AAB[illegible]	Sequence 1, Appl
43	100	[illegible]	AA[illegible]	Met Ala human growth h
44	100	[illegible]	AAB[illegible]	Pro Leu human growth h
45	100	[illegible]	AAB[illegible]	Leu Leu human growth h

ALIGNMENTS

RESULT 1
AAY12[illegible]
ID AAY12[illegible] standard; protein; 49 AA.
XX
AC AAY12[illegible];
XX
DT [illegible]
XX
DE Human growth hormone (hGH) N-terminal [illegible]
XX
KW [illegible]
KW Hordinar [illegible] chimeric protein [illegible]
KW pituitary [illegible]
XX
OS Homo sapiens.
XX
PN WO[illegible]
XX
PD [illegible]
XX
PF [illegible]
XX
PR [illegible]
XX
RA [illegible]
XX
RT Claim 2;
XX
RL WPI; [illegible]
XX
CC New chimeric protein [illegible] is claimed. The protein is [illegible]
CC pituitary [illegible] [illegible] [illegible] [illegible] [illegible]

PS Claim 4; Page 28; 46pp; English.
XX
CC This sequence represents an N-terminal fragment of human growth
CC hormone (hGH) which is a component of a chimeric protein,
CC hGH mini-proinsulin (AAY42857). The hGH portion of the chimeric protein
CC acts as an intramolecular chaperone (IMC) for the insulin precursor,
CC enabling it to fold correctly. A cleavable peptide linker with a
CC C-terminal Arg residue (AAY42857) enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC provide human insulin with correctly linked cysteine bridges with
CC fewer necessary procedural steps, and hence resulting in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intermolecular interactions among
CC the fusion proteins, thus allowing folding of the fused insulin precursor
CC at commercially useful high concentrations. The procedural steps of
CC cyanogen bromide cleavage, oxidative sulphitolysis and related
CC purification steps can thus be eliminated, also with the use of high
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 430 AA;

Query Match 100.0%; Score 2060; DB 20; Length 430;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 2
AAY42857
ID AAY42857 standard; protein; 72 AA.
XX
AC AAY42857;
XX
DT 19-JAN-2000 (first entry)
XX
DE hGH mini proinsulin chimeric protein.
XX
KW Insulin precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX
OS Synthetic.
OS Homo sapiens.
XX
RN [1]
XX
RP 1-72
RX DOI; 10.1016/S0378-1119(99)00562-2.
XX
RX PUBMED; 10606766.
XX
RA (INSU) INSULIN GENTECH LTD;
XX
RT "Gene 1";
XX
RL WPI; 2000-618869/57.
XX
PN New chimeric proteins containing human growth hormone fragment, used
PN particularly for the production of human insulin
XX
PS Claim 1; Page 30; 46pp; English.
XX
CC This sequence represents a chimeric protein, hGH mini proinsulin.
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42855, a cleavable peptide
CC linker (AAY42857), and a human insulin precursor comprising insulin
CC A and B chains (AAY42856). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor,
CC enabling it to fold correctly. The C-terminal Arg residue which enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC provide human insulin with correctly linked cysteine bridges with
CC fewer necessary procedural steps, and hence resulting in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intermolecular interactions among
CC the fusion proteins, thus allowing folding of the fused insulin precursor
CC at commercially useful high concentrations. The procedural steps of
CC cyanogen bromide cleavage, oxidative sulphitolysis and related
CC purification steps can thus be eliminated, also with the use of high
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 430 AA;

Query Match 100.0%; Score 2060; DB 20; Length 430;
Best Local Similarity 100.0%; Pred. No. 2e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 3
AAY42860
ID AAY42860 standard; protein; 107 AA.
XX
AC AAY42860;
XX
DT 19-JAN-2000 (first entry)
XX
DE hGH mini proinsulin chimeric protein.
XX
KW Insulin precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX
OS Synthetic.
OS Homo sapiens.
XX
RN [1]
XX
RP 1-107
RX DOI; 10.1016/S0378-1119(99)00562-2.
XX
RX PUBMED; 10606766.
XX
RA (INSU) INSULIN GENTECH LTD;
XX
RT "Gene 1";
XX
RL WPI; 2000-618869/57.
XX
PN New chimeric proteins containing human growth hormone fragment, used
PN particularly for the production of human insulin
XX
PS Claim 1; Page 30; 46pp; English.
XX
CC This sequence represents an N-terminal fragment of human growth hormone
CC (hGH) which is a component of a chimeric protein (AAY42861) which also
CC contains a human insulin precursor (AAY42860). The hGH portion of the
CC chimeric protein acts as an intramolecular chaperone (IMC) for the
CC insulin precursor, enabling it to fold correctly. A cleavable peptide
CC linker with a C-terminal Arg residue (AAY42857) enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC and related purification steps can thus be eliminated, also with the
CC use of high concentrations of mercaptan or the use of hydrophobic
CC absorbent resins.
XX
SQ Sequence 107 AA;

Query Match 100.0%; Score 2060; DB 20; Length 107;
Best Local Similarity 100.0%; Pred. No. 4.1e-25;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

PT provide human insulin with correctly linked cysteine bridges with
PT fewer necessary processing steps, and hence results in a higher yield
PT of human insulin. The IMP sequences not only protect insulin sequences
PT from intracellular degradation by a microorganism host, but also promote
PT the initial folding of the fused insulin precursor, facilitate the solubility of
PT the fusion protein and decrease the intermolecular interactions among
PT the fusion proteins, thus allowing folding of the fused insulin precursor
PT at commercially useful high concentrations. The procedure, steps of
PT cyanogen bromide cleavage, oxidative sulphitolysis and related
PT purification steps can thus be eliminated, along with the use of high
PT concentrations of mercaptan or the use of hydrophobic adsorbent resins.
XX
SQ Sequence 167 AA;

Query Match 100.0%; Score 26; DB 2; Length 167;
Best Local Similarity 100.0%; Pred. No. 4.8e-25;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MEPICPLSPIFINAMLRAHRLH 20
|||||||||||||||||||||
Db 1 MEPICPLSPIFINAMLRAHRLH 20

RESULT 4
AAW[illegible]
ID AAW[illegible] standard; protein; 144 AA.
XX
AC AAW[illegible];
XX
DT [illegible] (first entry)
XX
DE human anti-angiogenic peptide 16K 67B Met Tr[illegible].
XX
KW Human anti-angiogenic peptide; placental [illegible]; anti-angiogenesis;
KW [illegible] with hormone; hGH; hGH V; capillary endothelial cell proliferation;
KW placental vascularisation; premature; treatment; angiogenic disease;
KW [illegible] inhibitor; malignant; [illegible]; arteriovenous malformation;
KW arthritis; atherosclerotic plaque; corneal graft neovascularisation;
KW [illegible] and hepatic proliferative retinopathy; macular degeneration; trachoma;
KW [illegible]; [illegible]; retrolental [illegible]; sickle cell anaemia;
KW psoriasis; [illegible]; [illegible]; Kaposi's sarcoma; [illegible];
KW [illegible]; [illegible]; [illegible]; fracture; [illegible] Weber syndrome;
KW [illegible]; pre-eclampsia; intrauterine growth retardation;
KW [illegible] dysfunction.
XX
OS Homo sapiens.
XX
PN W[illegible]-A.
XX
PD 14-NOV-[illegible].
XX
PF 14-MAY-[illegible] 96WO-US06[illegible].
XX
PR 14-MAY-[illegible] 95US-[illegible].
XX
PA ([illegible]) UNIV CALIFORNIA.
XX
PI Martin GA, Strauss EJ, Taylor SD, Wechsler FS;
XX
DR WPI; [illegible]-[illegible].
DR N-PSDB; AAX [illegible].
XX
PT New anti-angiogenic peptides - comprise N-terminal fragments of
PT human placental lactogen, human growth hormone or variant
PT variant or human prolactin
XX
SF [illegible] to [illegible] - [illegible] - [illegible].
XX
CC This invention describes novel human anti-angiogenic peptides derived
CC from the N-terminal end of naturally occurring mature human placental
CC lactogen (hPL), human growth hormone (hGH), growth
CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
CC capillary endothelial cell proliferation and translation, (ii) inhibit

CC metastatic [illegible] and [illegible]
CC [illegible] and [illegible] [illegible]
CC hGH, the preferred [illegible] hGH V, the [illegible] [illegible]
CC [illegible] a probable [illegible], [illegible] [illegible]
CC premature, the peptides can be used [illegible] [illegible]
CC a subject. It [illegible] can be used [illegible]
CC modulating vascularisation of a patient's placenta, in particular, [illegible]
CC peptides can be used [illegible] [illegible]
CC angiogenic [illegible] [illegible] [illegible]
CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC delayed wound healing, pr[illegible] tumour growth, such as [illegible]
CC retinopathy, macular degeneration, trachoma, [illegible] [illegible]
CC in haemophilic joints, inappropriate vascularisation in wounded tissues
CC such as hypertrophic scars or keloid scars, neovascularisation [illegible]
CC tumour, uveitis, [illegible], [illegible] Weber syndrome, [illegible]
CC pyogenic granuloma, retrolental fibroplasia, scleroderma, [illegible]
CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic [illegible],
CC [illegible] and [illegible] [illegible] [illegible] [illegible]
CC and [illegible] tissue. They can also be used as contraceptive agents. DNA
CC encoding the peptides can be used in gene therapy. The [illegible]
CC arterial [illegible] [illegible] [illegible], hGH, hGH-V, [illegible]
CC can be used in assays for [illegible] [illegible] [illegible]
CC with pre-eclampsia, intrauterine growth retardation and [illegible]
CC dysfunction.
XX
SQ Sequence 144 AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches 4; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 5
AAI[illegible]
ID AAI[illegible] standard; protein; 61 AA.
XX
AC AAI[illegible];
XX
DT [illegible] (first entry)
XX
DE Human growth hormone mutant.
XX
KW Human growth hormone; [illegible] [illegible];
KW [illegible] [illegible] [illegible] [illegible] [illegible] [illegible];
XX
OS Homo sapiens (human).
XX
PN EP-[illegible]-A.
XX
PD [illegible]-JUL-[illegible].
XX
PF [illegible]-JAN-[illegible] 90EP-[illegible].
XX
PR [illegible]-JAN-[illegible] 89US-[illegible].
XX
PA (GENE) GENENTECH INC.
XX
PI [illegible] DA;
XX
DR WPI; [illegible]-[illegible].
XX
PT New human growth hormone variants - have one or more [illegible]
PT [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
PT deletion, etc.
XX
PS [illegible] pair [illegible].
XX
CC Human growth hormone variant [illegible] [illegible] [illegible]
CC protein which [illegible] [illegible] [illegible] [illegible]

CC growth factor (beta-NGF) at the C-terminal. Beta NGF can be used to
CC control geriatric dementia and other nervous disorders, and can be
CC released from the fusion protein by incubation with thrombin (see
CC AAN[illegible], AAP[illegible], AAP[illegible]).
XX
SQ Sequence 149 AA;

Query Match 100.0%; Score 260; DB 10; Length 149;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 6
AAY4[illegible]
ID AAY4[illegible] standard; protein; 150 AA.
XX
AC AAY4[illegible];
XX
DT [illegible]-JAN-2000 (first entry)
XX
DE Chimeric protein. SEQ ID 7.
XX
KW Insulin; precursor; growth hormone; chaperone; intramolecular;
KW folding; conformation; chimeric protein; cleavable; recombinant;
KW production; yield.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO[illegible] A1.
XX
PD 07-[illegible]-1999.
XX
PF 31-MAR-1998; 98WO-CN00052.
XX
PR 31-MAR-1998; 98WO-CN00052.
XX
PA (TSNG-) TSINGHUA GANIEN BIOTECHNOLOGY LTD.
XX
PI [illegible]
XX
DR WPI; 1999-[illegible]/52.
XX
PT New chimeric proteins containing human growth hormone fragment, used
PT particularly for the production of human insulin -
XX
PS Claim 14; Page 30-31; 46pp; English.
XX
CC This sequence represents a chimeric protein, which contains an
CC N-terminal fragment of human growth hormone (hGH) of the sequence given
CC in AAY42856, a cleavable peptide linker (AAY42857), and a human insulin
CC precursor comprising insulin A and B chains (AAY42859). The hGH portion
CC of the chimeric protein acts as an intramolecular chaperone (IMC) for
CC the insulin precursor, enabling it to fold correctly. The cleavable
CC peptide linker has a C-terminal Arg residue which enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence resulting in a higher
CC yield of human insulin. The IMC sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursor, facilitate the
CC solubility of the fusion protein and decrease the intermolecular
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC procedural steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC and related purification steps can thus be eliminated, along with the use
CC of high concentrations of mercaptan or the use of hydrophobic absorbent
CC resins.
XX
SQ Sequence 150 AA;

Query Match [illegible]; Score [illegible]; DB 20; Length [illegible];
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 7
AAP90129
ID AAP90129 standard; protein; 192 AA.
XX
AC AAP90129;
XX
DT 06-FEB-1996 (revised)
DT 03-NOV-1989 (first entry)
XX
DE Human growth hormone.
XX
KW Human growth hormone; fusion protein; recombinant
KW vector.
XX
OS Homo sapiens (Human).
XX
PN JP01144981 A.
XX
PD 07-JUN-1989.
XX
PF 02-DEC-1987; 87JP-[illegible]1987.
XX
PR 02-DEC-1987; 87JP-[illegible]1987.
XX
PA (WAKU) WAKUNAGA SEIYAKU KK.
XX
DR WPI; 1989-203384/29.
DR N-PSDB; AAN90269.
XX
PT Recombinant vector contg. fusion protein - consisting of human
PT growth hormone or deriv. ligated to foreign protein, for stability
PT and high yield.
XX
PS Disclosure; Fig 1; [illegible]pp; Japanese.
XX
CC The invention consists of a vector contg. a fusion protein which is
CC formed by ligating, downstream of a promoter, hGH or a deriv. (deriv.
CC formed by substn. of Met-14 with Leu) and a foreign protein.
CC Stability of the vector in the host is greatly increased so the
CC protein yield is higher.
XX
SQ Sequence 192 AA;

Query Match 100.0%; Score 260; DB 10; Length 192;
Best Local Similarity 100.0%; Pred. No. [illegible];
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
|||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 8
AAW92264
ID AAW92264 standard; protein; 192 AA.
XX
AC AAW92264;
XX
DT 08-JUN-1999 (first entry)
XX
DE Human [illegible] signal peptide hGH Met [illegible].

XX
KW Human anti-angiogenic protein; placenta; anti-HCG; anti-hCG;
KW anti-hCG beta; anti-hCG V; capillary endothelial; pre-eclamptic;
KW placental vascularisation; pregnancy; treatment; angiogenic disease;
KW tumour; infarct; malignant; anti-fibrous; arteriovenous malformation;
KW arthritis; atherosclerotic plaque; corneal graft neovascularisation;
KW wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW granulation; glaucoma; ocular uveitis; fracture; osler-Weber syndrome;
KW psoriasis; tie protein; selerderma; Kaposi's sarcoma; vascular adhesion;
KW ovary; luteinized; reproductive disorder; contraceptive agent;
KW gene therapy; pre-eclampsia; intrauterine growth retardation;
KW placental dysfunction.
XX
OS Homo sapiens.
XX
PN WO9954455-A1.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1999; 99WO-GB001151.
XX
PR 16-APR-1998; 98GB-0007972.
XX
PA (MEDI-) UNIV WALES COLLEGE MEDICINE.
XX
IN Marshall JM; Stewart IJ; Taylor PD; Weller RO;
XX
DR WPI; 1999-624176/53.
DR N-PSDB; AAX17744.
XX
PT New anti-anti-hCG peptides comprise N-terminal fragments of
PT human placental lactogen, human or with hCG beta or with hCG
PT variant or human prolactin.
XX
PS Example 6; Page 15; 27pp; English.
XX
CC This invention describes novel human anti-angiogenic peptides derived
CC from up to 100 consecutive amino acids selected from the N-terminal end
CC of human placental lactogen (hPL), human or with hCG beta (hCG beta)
CC hormone variant (hCG V), or human prolactin. Such peptides (i) inhibit
CC capillary endothelial cell proliferation and organisation (ii) inhibit
CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at
CC least one specific receptor which does not bind an intact full length
CC hCG, hPL, prolactin or hCG-V. The invention also describes a method for
CC diagnosing a prognostic abnormality of placental vascularisation during
CC pregnancy. The peptides can be used for treating an angiogenic disease in
CC a subject, for inhibiting tumour formation or growth in a patient or for
CC modulating vascularisation of a patient's placenta. In particular, the
CC peptides can be used for preventing or treating e.g. malignant tumours,
CC arteriovenous malformation, arthritis such as rheumatoid
CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC haemophilic joint, wound healing, proliferative retinopathy such as diabetic
CC retinopathy, macular degeneration, granulation such as those occurring
CC in haemophilic joints, inappropriate vascularisation in wound healing
CC such as hypertrophic scars or keloid scars, hemangioma, chronic ulcers,
CC ocular uveitis, fracture, atherosclerosis, psoriasis,
CC pyogenic granuloma, retrolental fibroplasia, scleroderma, solid tumours,
CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic cervicitis,
CC endometriosis, and reproductive disorders such as follicular and luteal cysts
CC and hydatidiform mole. They can also be used as contraceptive agents. DNA
CC encoding the peptides can be used in gene therapy. The measurement of
CC anti-hCG beta or N-terminal fragments of hCG, hCG V, prolactin or hPL
CC can be used in assays for impairment of fetal development associated
CC with pre-eclampsia, intrauterine growth retardation and placental
CC dysfunction.
XX
SQ Sequence 100 AA;

Query Match 100.0%; Score 100; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 9.1e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPFPLPSLFSNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MPFPLPSLFSNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50

RESULT 1
AAP01255
ID AAP01255 standard; protein; 211 AA.
XX
AC AAP01255;
XX
DT 24-DEC-1998 (first entry).
XX
DE Human beta-hCG with last 3 additional amino acids for use in pregnancy
XX
KW Human beta-hCG with last 3 C-terminal amino acids; pregnancy;
KW detection; beta-hCG; human chorionic gonadotrophin; hCG;
XX
OS Homo sapiens (human).
XX
FH Key Location/Qualifiers
FT Peptide 1..211
FT Peptide 1..211
FT Peptide 1..211
XX
PN EP-A-0849283-A1.
XX
PD 24-JUN-1998.
XX
PF 17-DEC-1997; 97EP-0310322.
XX
PR 20-DEC-1996; 96GB-0026577.
XX
PA (SCLA-) SCLAVO SPA.
XX
PI Bellisario R;
XX
DR WPI; 1998-347176/30.
XX
PT New human beta-hCG with last 3 amino acids additional to pregnancy
PT testing "kit" for the early detection of pregnancy, useful for treating
PT dementia, etc.
XX
PS Claim 10; Page 17; 22pp; English.
XX
CC Fusion protein consisting of human beta-hCG subunit and the
CC N-terminal end (1st peptide) of beta-hCG subunit (2nd peptide) is
CC thrombin recognition sites and human beta-nerve growth factor (beta-NGF)
CC at the C-terminus. Beta-NGF can be used for treating peripheral
CC and other neurological disorders and can be produced in the fusion
CC protein by incubation with thrombin (see AAN01256 and AAP01254).
XX
SQ Sequence 211 AA;

Query Match 100.0%; Score 100; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPFPLPSLFSNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50
|||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MPFPLPSLFSNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 50

RESULT 10
AAP01254
ID AAP01254 standard; protein; 211 AA.
XX
AC AAP01254;
XX
DT 24-DEC-1998 (first entry).
XX
DE Human beta-hCG with last 3 C-terminal amino acids
XX

```
KW   beta NGF; E.coli; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         145..262
XX
PN   JP61205485-A.
XX
PD   11-SEP-1986.
XX
PF   08-MAR-1985;   85JP-004577.
XX
PR   08-MAR-1985;   85JP-004577.
XX
PA   (HISZ) HISHIKAWA [illegible]
XX
DR   WPI; 1986-281646/43.
XX
PT   Gene segment of human nerve growth factor - used in prodn. of
PT   NGF producing recombinant Escherichia strain.
XX
PS   Claim 32; Page 482; 71pp; Japanese.
XX
CC   The protein is a direct translation of the upstream tryptophan
CC   promoter operator lacking it's attenuation sequence and human
CC   beta NGF sequence. The product may be efficiently expressed from a
CC   transformed E.coli expression system.
CC   See also AAN60816-7.
XX
SQ   Sequence   262 AA;

  Query Match             100.0%;  Score 260;  DB 7;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49


RESULT 11
AAR11740
ID   AAR11740 standard; Protein; 262 AA.
XX
AC   AAR11740;
XX
DT   25-JUN-1991  (first entry)
XX
DE   Human growth hormone/human nerve growth factor beta fusion protein.
XX
KW   hGH; hNGF; nervous system diseases; dementia.
XX
OS   Homo sapiens.
XX
PN   JP[illegible]-A.
XX
PD   12-MAR-1991.
XX
PF   07-AUG-1989;   89JP-0202845.
XX
PR   07-AUG-1989;   89JP-0202845.
XX
PA   (SUMO) SUMITOMO [illegible].
XX
DR   WPI; 1991-128768/18.
DR   N-PSDB; AAQ11578.
XX
PT   Purificn. of human neuron growth factor beta subunit-contg. protein
PT   by contacting with gel having cation exchange gp. in presence of
PT   urea
XX
PS   Disclosure; Fig 1; 7pp; Japanese.
```

```
XX
CC   A recombinant human nerve growth factor beta subunit containing protein
CC   can be prepared as this fusion protein. It is purified by contacting
CC   a gel having a cation exchange gp. with the fusion protein, in the
CC   presence of urea. The purified protein is useful in a medicament
CC   for treating disorders of the nervous system, eg dementia.
XX
SQ   Sequence   262 AA;

  Query Match             100.0%;  Score 260;  DB 12;  Length 262;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49


RESULT 12
AAR03255
ID   AAR03255 standard; protein; 315 AA.
XX
AC   AAR03255;
XX
DT   19-JUL-1990  (first entry)
XX
DE   Fusion protein of B cell stimulatory factor 2 and B cell
DE   differentiation factor.
XX
KW   B cell stimulatory factor 2; interleukin 6; B cell differentiation;
KW   interleukin 6; fusion protein.
XX
OS   Homo sapiens.
XX
PN   JP02[illegible]-A.
XX
PD   17-JAN-1990.
XX
PF   01-JUL-1988;   88JP-0162556.
XX
PR   01-JUL-1988;   88JP-0162556.
XX
PA   (I YT ) [illegible]
XX
DR   WPI; 1990-062207/09.
DR   N-PSDB; AAQ[illegible].
XX
PT   Prepn. of human B cell differentiation factor - from specified DNA
PT   sequence segment, by recombinant DNA techniques, allows prodn. of
PT   specified amino acid sequence.
XX
PS   Claim 1; Page [illegible]; [illegible]pp; Japanese.
XX
CC   The protein is produced by fusing DNA encoding BCDF (IL-6) with DNA
CC   encoding BSF-2 (IL-6) and inserting the product into an expression vector.
CC   See also AAR05311 and AAR05313.
XX
SQ   Sequence   315 AA;

  Query Match             [illegible];  Score [illegible];  DB 11;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. 1.3e-24;
  Matches   49;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MEFIIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49


RESULT 13
AAR05313
ID   AAR05313 standard; protein; 141 AA.
XX
AC   AAR05313;
```

XX
ID [illegible] (first entry)
XX
DE Sequence of B-cell stimulatory factor-2 (BSF-2).
XX
KW B-cell stimulatory factor-2; interleukin-6.
XX
OS [illegible] sapiens.
XX
PN [illegible] A.
XX
PD [illegible]-JAN-1990.
XX
PF [illegible]-1990; [illegible]-JAN-1990.
XX
PR [illegible]-1989; [illegible]-JAN-1990.
XX
RA [illegible] T.S.B.M.Sc.
XX
RT "[illegible]";
RL N.[illegible] AA; [illegible].
XX
CC Single, 28 kDa B-cell differentiation factor [illegible] specified DNA
CC sequence segment, by recombinant DNA technique, gives protein of
CC specified amino acid sequence.
XX
CC Publication note: [illegible] Japanese.
XX
CC Recombinant protein of this protein can be fused with [illegible] DNA encoding B-cell
CC differentiation factor (BSF-2, and ligated into an expression vector to
CC produce [illegible] fusion protein.
CC See also AAB[illegible].
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; protein; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible] (first entry)
XX
DE Human growth hormone with collagenase recognition site.
XX
KW Human; [illegible] with collagenase recognition site.
XX
OS Homo sapiens.
XX
PN [illegible].
XX
PD [illegible]-MAY-[illegible].
XX
PF [illegible]-[illegible]-[illegible].
XX
PR [illegible]-[illegible]-[illegible].
XX
RA [illegible] [illegible] [illegible].
XX
RT "[illegible]";
RL [illegible] [illegible] [illegible].
XX
DR [illegible]; [illegible].
XX
FH Key Location/Qualifiers
FT [illegible]
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; protein; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible] (first entry)
XX
DE Sequence of protein with human mature interleukin-6.
XX
KW Protein; human.
XX
OS Synthetic.
XX
PN [illegible] A.
XX
PD [illegible]-DEC-1990.
XX
PF [illegible]-[illegible]-[illegible].
XX
PR [illegible]-[illegible]-[illegible].
XX
RA [illegible] T.S.; [illegible] M.
XX
DR [illegible]; [illegible].
XX
FH Key Location/Qualifiers
FT Polypeptide with [illegible] [illegible].
FT [illegible] [illegible] [illegible] [illegible] [illegible].
FT Segment with specified DNA sequence
XX
PS Claim [illegible]; [illegible] [illegible].
XX
CC The polypeptide (AAB[illegible]) with [illegible] [illegible] [illegible] [illegible]
CC (AAN81005) encoding it are claimed. A DNA construct [illegible] [illegible]
CC the N-terminus. The polypeptide [illegible] [illegible] [illegible] [illegible]
CC including cancer. [illegible] [illegible] [illegible] [illegible] [illegible]
CC production rate and is useful [illegible] [illegible] [illegible] [illegible]
CC and refined for use as a [illegible] [illegible] [illegible].
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]

Db [illegible]

Search completed: [illegible], [illegible] seconds
Job time : [illegible] secs

Run on: February 5, 2003, 12:14:21 ; Search time 1.08 Seconds
(without alignments)
[illegible] Million Cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQ 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 28124 seqs, 9124412 residues

Total number of hits satisfying chosen parameters: 28124

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

The N... is the number of results probabilistically shorter than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB	ID	Description
1	245	94.2	217	1	SCHU	somatotropin I pre
2	238	91.5	217	2	S[illegible]	somatotropin - the
3	228	87.7	217	1	SIHUV	somatotropin 2 pre
4	228	87.7	217	1	SIHUV2	somatotropin 2 pre
5	225	86.5	222	2	[illegible] somatotro
6	218	83.8	217	2	I5[illegible]	chorionic somatomam
7	216	83.1	217	1	[illegible]	somatotropin - the
8	201	77.3	217	2	I67409	chorionic somatomam
9	197	75.8	217	2	A[illegible]	chorionic somatotropin
10	197	75.8	217	1	I67402	chorionic somatotropin
11	197	75.8	217	2	E[illegible]	chorionic somatotropin
12	[illegible]	[illegible]	22	1	S[illegible]	somatotropin precu
13	[illegible]	[illegible]	190	2	PN 14[illegible]	somatotropin - sei
14	[illegible]	[illegible]	24	2	S[illegible]	somatotropin - Ba
15	[illegible]	[illegible]	24	2	[illegible]	somatotropin - Ath
16	[illegible]	[illegible]	24	2	JH[illegible]	somatotropin - Ath
17	[illegible]	[illegible]	22	2	S[illegible]	somatotropin precu
18	[illegible]	[illegible]	20	1	SIPJ	somatotropin precu
19	[illegible]	[illegible]	22	2	[illegible]	somatotropin precu
20	[illegible]	[illegible]	216	2	S49[illegible]	somatotropin precu
21	[illegible]	[illegible]	22	2	[illegible]	somatotropin - sei
22	[illegible]	[illegible]	22	2	JH[illegible]	somatotropin precu
23	[illegible]	[illegible]	216	2	A[illegible]	somatotropin precu
24	[illegible]	[illegible]	190	2	AC[illegible]	somatotropin - sei
25	156	57.7	216	2	JC[illegible]	somatotropin precu
26	142	[illegible]	191	2	A[illegible]	somatotropin - sei
27	140	[illegible]	190	2	JN[illegible]	somatotropin - sei
28	134	49.4	190	2	S21[illegible]	somatotropin - Rhe
29	134	49.4	216	2	A[illegible]	somatotropin precu
30	134	49.4	217	1	S[illegible]	somatotropin precu
31	134	49.4	217	1	S[illegible]	somatotropin precu
32	134	49.4	217	1	S[illegible]	somatotropin precu
33	134	49.4	217	1	S[illegible]	somatotropin precu
34	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
35	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
36	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
37	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
38	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
39	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
40	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
41	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
42	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
43	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
44	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu
45	13[illegible]	49.0	217	1	S[illegible]	somatotropin precu

ALIGNMENTS

RESULT 1
SCHU
Somatotropin I precursor - human
N:Alternate names: growth hormone; GH; pituitary growth hormone
C:Species: human (primates)
C:Date: 24-Apr-1991 #sequence_revision 24-Apr-1991 #text_change 09-Jul-2002
C:Accession: A[illegible]; A[illegible]; A[illegible]; A[illegible]; A[illegible]; A[illegible]; A[illegible]; A[illegible]
R:DeNoto, F.M.; Moore, D.D.; Goodman, H.M.
Nucleic Acids Res. [illegible]
A:Title: Human growth hormone cDNA sequence and expression in non-pituitary tissues
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Accession: A[illegible]
A:Molecule type: mRNA
A:Residues: 1-217 <DEN
A:Cross-references: GB:V[illegible]
A:Notes: the 20K short form somatotropin lacks residues 50-[illegible] (q.v.); it is the [illegible]
R:Cohen, F.E.; Riek, R.P.; Sosnick, T.R.; Prusiner, S.B.; Baldwin, R.A.; Kuntz, I.D.
Genomics 4, 479-497, 1989
A:Title: The human growth hormone gene family: Structure and evolution of the sequences [illegible]
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Accession: A[illegible]
A:Molecule type: mRNA
A:Residues: 1-217 <Coh
A:Cross-references: GB:V[illegible]; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
R:Keshet, Y.; Seidman, J.G.
Nucleic Acids Res. [illegible]
A:Title: Multiple structural variants of the human growth hormone gene
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Accession: A[illegible]
A:Molecule type: mRNA
A:Residues: 1-217 <Kes
A:Cross-references: GB:V[illegible]
A:Notes: 20 Pro was the 1 and
R:Martial, J.A.; Baxter, J.D.; Goodman, H.M.; Seeburg, P.H.
Science 205, 602-607, 1979
A:Title: DNA sequences coding for human growth hormone precursors
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Accession: A[illegible]
A:Molecule type: mRNA
A:Residues: 1-217 <Mar
R:[illegible]; [illegible]; [illegible]; [illegible]; [illegible]
Arch. Biochem. Biophys. [illegible]
A:Title: Human pituitary growth hormone. XXX. The primary structure of the 20,000-dalton variant
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Contents: annotation
R:[illegible]; [illegible]; [illegible]
Arch. Biochem. Biophys. [illegible]
A:Title: Human pituitary growth hormone. XXX. The primary structure of the 20,000-dalton variant
A:Reference number: A[illegible]; MUID:[illegible]; PMID:[illegible]
A:Accession: A[illegible]
A:Molecule type: protein

A;Residues: 27-94;96-217 <LI>
E;Nicoll, R.E.
Nature New Biol., 245, 204-205, 1973
A;Title: Revised primary structure for human growth hormone.
A;Reference number: A93397; MUID:73154268; PMID:4577646
A;Accession: A93397
A;Molecule type: protein
A;Residues: 27-191 <NIA>
E;Nicoll, R.E.; Botelho, M.C.; Sasaki, F.; Rosenblum, I.Y.; Greenwood, F.C.
Proc. Natl. Acad. Sci. U.S.A., 68, 866-869, 1971
A;Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution
A;Reference number: A93770; MUID:71153964; PMID:5278439
A;Accession: A93770
A;Molecule type: protein
A;Residues: 110-120;157-159 <NIZ>
E;Nicoll, R.E.
In Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop, Griffiths, K.,
A;Title: The chemistry of the human lactogenic hormones.
A;Reference number: A94427
A;Contents: annotation; somatotropin revision
E;Li, C.H.; Dixon, J.S.; Li, C.H.
Int. J. Pept. Protein Res. 4, 281-287, 1972
A;Title: Sequence comparison of human pituitary growth hormone, human chorionic somatom
A;Reference number: A91764; MUID:74092028; PMID:4675454
A;Accession: A91764
A;Molecule type: protein
A;Residues: 27-217 <BEW>
E;Chen, H.; Li, C.H.
Biochem. Biophys. Res. Commun., 92, 511-516, 1980
A;Title: The 2'-deamidation [87-Asn] of human growth hormone I. Isolation of the mono-
A;Reference number: A90217; MUID:80179796; PMID:6767439
A;Contents: somatotropin, 20K short variant
A;Accession: A90217
A;Molecule type: protein
A;Residues: 1-57;75-80 <LEW>
E;Chapman, G.E.; Rogers, K.M.; Brittain, L.; Bradshaw, R.A.; Bates, G.D.; Turner, J.
J. Biol. Chem., 256, 2395-2401, 1981
A;Title: The 20,000 molecular weight variant of human growth hormone. Preparation and
A;Reference number: A92411; MUID:81117661; PMID:7462247
A;Contents: somatotropin, 20K short variant
A;Accession: A92411
A;Molecule type: protein
A;Residues: 27-57;59-79 <NIA>
E;Sloan, R.N.P.; Searcy, R.E.; Levis, L.J., 3 mls, U.S.
J. Protein Chem., 2, 425-436, 1983
A;Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
A;Reference number: A61466
A;Accession: A61466
A;Molecule type: protein
A;Residues: 27-64 <SIN>
A;Note: growth hormone 5K peptide has insulin potentiating activity; its physiological
E;Bewley, T.A.; Li, C.H.
Biol. Chem. Hoppe Seyler, 371, 433-441, 1990
A;Title: Identification of the aspartimide structure in a previously reported peptide
A;Reference number: S00868; MUID:90384745; PMID:2378679
A;Accession: S00868
A;Molecule type: protein
A;Residues: 27-44,'L',46-47 <FOR>
E;de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.
Science, 255, 306-312, 1992
A;Title: Human growth hormone and extracellular domain of its receptor: crystal struct
A;Reference number: A41728; MUID:92196022; PMID:1549776
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
A;Note: the structure of the complex with growth hormone receptor is described
E;Cohen, F.E.; Baldridge, J.S.; McKeown, K.S.; Reynaker, H.L.; Chang, C.N.
Science, 247, 954, 1990
A;Title: Periplasmic production of correctly processed human growth hormone in Escheri
A;Reference number: I41127; MUID:90615709; PMID:2172261
A;Accession: I04549
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-26 <BES>
A;Cross-references: GB:M13438; NID:g181455; PIDN:AAA52760.1; PID:g181456

C;Comment: The gene for this hormone is transcribed only in somatotropic cells. Th
C;Comment: About 99% of somatotropic pit is the 22K and 1 20.
C;Genetics:
A;Gene: GDB:GH1
A;Cross-reference: GDB:120290; OMIM:139250
A;Map position: 17q22-17q24.3
A;Introns: 1/21; 52/3; 3/4; 153/4
C;Superfamily: prolactin
C;Keywords: alternative splicing; hormone; pituitary
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-217/Product: somatotropin 1, 22K form #status experimental <PRO>
F;27-69/Misc-difference: growth hormone 5K peptide #status experimental <5KP>
F;27-57,73-217/Product: somatotropin 1, short form #status experimental <SHO>
F;79-191,206-217/Disulfide bonds: #status experimental

Query Match 99.1%; Score 215; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||
Db 169 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 218

RESULT 2
167410
somatotropin - rhesus monkey
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus monkey)
C;Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 16-Jul-1998
C;Accession: I67410; A05044
R;Golos, T.G.; Herring, M.; Fisher, J.M.; Fowler, P.B.
Endocrinology 133, 1744-1752, 1993
A;Title: Cloning of four growth hormone/chorionic somatomammotropin genes from
A;Reference number: I54267; MUID:94008724; PMID:8104617
A;Accession: I67410
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-217 <RES>
A;Cross-references: GB:L19556; NID:g396111; PIDN:AAA19842.1; PID:g396112
R;Li, C.H.; Chung, D.; Lahm, H.W.; Stein, S.
Arch. Biochem. Biophys. 245, 287-291, 1986
A;Title: The primary structure of monkey pituitary growth hormone
A;Reference number: A25994; MUID:86172455; PMID:3082659
A;Accession: A25994
A;Molecule type: protein
A;Residues: 27-44,'L',46-47,'E',49-64,'E',66-67
A;Note: the monkey species is not identified in the reference
E;Raben, M.S.
Science 125, 883-884, 1957
A;Title: Preparation of growth hormone from pituitaries of non-human
A;Reference number: A14771
A;Contents: annotation; identification of source organism
C;Superfamily: prolactin

Query Match 99.1%; Score 215; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
||||||||||||||||||||||||||||||||||||||||||||||||
Db 169 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 218

RESULT 3
S18DUV
somatotropin 2 precursor - human
N;Alternate names: growth hormone 2; growth hormone variant; hGH-V; placental somatot
N;Contains: somatotropin 2; Lmw species; somatotropin 2, short [up to 191]
C;Species: Homo sapiens (man)
C;Date: 12-Dec-1986 #sequence_revision 16-Jul-1998 #text_change 09-Jul-1999
C;Accession: S18DUV; P01242; A01711; A93761; A97761
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Seeburg, P.H.; Goodman

RESULT 7
I67411
source Triple Macaque
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31 May 1996 #sequence_revision 31 May 1996 #text_change 16 Jul 1999
C;Accession: I67411
R;Bolos, L.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744 1752, 1993
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementar
A;Reference number: 153267; MUID:94008724; PMID:8404617
A;Accession: I67411
A;Status: preliminary; translated from GB/EMBL/GenBank
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:L17154; NID:g297122; PIDN:AAA19841.1; PID:g297123
C;Superfamily: prolactin

Query Match 78.8%; Score 205; DB 2; Length 217;
Best Local Similarity 79.2%; Pred. No. 3.6e-38;
Matches 57; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MATPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 64
MA+PSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE
Db 1 MAIPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 60

RESULT 8
I67409
chorionic somatomammotropin-5 - rhesus macaque
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31 May 1996 #sequence_revision 31 May 1996 #text_change 16-Jul 1999
C;Accession: I67409
R;Bolos, L.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.
Endocrinology 133, 1744 1752, 1993
A;Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementar
A;Reference number: 153267; MUID:94008724; PMID:8404617
A;Accession: I67409
A;Status: preliminary; translated from GB/EMBL/GenBank
A;Molecule type: mRNA
A;Residues: 1-217 <BOL>
A;Cross-references: GB:L17154; NID:g297122; PIDN:AAA19841.1; PID:g297123
C;Superfamily: prolactin

Query Match 77.6%; Score 201; DB 2; Length 217;
Best Local Similarity 74.5%; Pred. No. 1.1e-37;
Matches 47; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MATPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 64
MA+PSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE
Db 1 MAIPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 60

RESULT 9
A26449
Chorionic somatomammotropin precursor (allele hCS-5) - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun 1988 #sequence_revision 30-Jun 1988 #text_change 28-Jul 1995
C;Accession: A26449
R;Hirt, H.; Kleiman, J.; Birnbaum, M.J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.; Ba
DNA 6, 59-70, 1987
A;Title: The human growth hormone gene locus: structure, evolution, and allelic variati
A;Reference number: A26449; MUID:87161235; PMID:3035480
A;Accession: A26449
A;Molecule type: DNA
A;Residues: 1-217 <HIR>
C;Superfamily: prolactin
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-217/Product: chorionic somatomammotropin hCS-5 variant #status predicted <MAT>

Query Match 75.8%; Score 197; DB 2; Length 217;
Best Local Similarity 79.2%; Pred. No. 1.7e-36;
Matches 57; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

Qy 1 MATPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 64
MA+PSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE
Db 1 MAIPSKLFLSAVLMVSLLIHFCQLAACPIPQTYGDDEKSIQPLGVQRLQEDVSRLRKDSIDIMKSRLQNEE 60

RESULT 10
LCHUG
chorionic somatomammotropin A precursor [validated] - human
N;Alternate names: chorionic somatomammotropin pit; placental lactogen
C;Species: Homo sapiens (man)
C;Date: 25-Oct 1985 #sequence_revision 25-Oct 1985 #text_change 16-Jul
C;Accession: A24345; A24346; A24347; A24348; A24349; A24350
R;Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera Saldana, H.A.; Gelinas, R.E.; Seeburg,
Genomics 4, 479-497, 1989
A;Title: The human growth hormone gene family: structures and evolutions
A;Reference number: A24345; MUID:89314272; PMID:2744260
A;Accession: A24345
A;Molecule type: DNA
A;Residues: 1-217 <CHE>
A;Cross-references: GB:M13438; NID:g178778; PIDN:AAA51523.1; PID:g178779
R;Goodman, H.M.; DeNoto, F.; Fiddes, J.C.; Hallewell, R.A.; Lee, J.; Moore, H.W.;
In Molecular and Possibly Genetic Information, 5001, W.A. Bentley Publ.
A;Reference number: A94422
A;Accession: A94422
A;Molecule type: mRNA
A;Residues: 1-217 <GOO>
R;Tanaka, M.; Maeda, N.; Watanabe, M.; Yamamoto, M.; Shimizu, K.; Nishi, T.; Nagashi
Biochem. Int. 7, 287-292, 1983
A;Title: cDNA cloning of human chorionic somatomammotropin I mRNA and its transcripts
A;Reference number: I92342; MUID:83209096; PMID:6835050
A;Accession: I92342
A;Status: translated from GB/EMBL/GenBank
A;Molecule type: mRNA
A;Residues: 1-8, 'AN>
A;Cross-references: GB:M14414; NID:g178778
R;Sherwood, L.M.; Bursztein, Y.; Schoettler, I.
Proc. Natl. Acad. Sci. U.S.A. 76, 6826-6830, 1979
A;Title: Primary structure of the NH2-terminal extra piece of the precursor of human
A;Reference number: A93855; MUID:80084267; PMID:291015
A;Accession: A93855
A;Molecule type: protein
A;Residues: 1-8, 'AN>
A;Experimental source: placenta
R;Shine, J.; Seeburg, P.H.; Martial, J.A.; Baxter, J.D.; Goodman, H.M.
Nature 270, 494-499, 1977
A;Title: Construction and analysis of recombinant DNA for human chorionic somatomam
A;Reference number: A94102; MUID:78051761; PMID:593608
A;Accession: A94102
A;Molecule type: DNA
A;Residues: 50-217 <SHI>
A;Experimental source: placenta
R;Li, C.H.; Dixon, J.S.; Lemaire, ...
Arch. Biochem. Biophys. 155, 95-110, 1973
A;Title: Amino acid sequence of human chorionic somatomammotropin
A;Reference number: A90154; MUID:74111721; PMID:4121437
A;Accession: A90154
A;Molecule type: protein
A;Residues: 27-217 <LI>
A;Experimental source: placenta
R;Niall, H.D.
In Endocrinology and Carbohydrate, 1975, E. Diczfalusy & Walker (eds.), pp. 1-12
A;Title: The chemistry of the human fetal-gonal hormones.
A;Reference number: A94427
A;Accession: A94427
A;Molecule type: protein
A;Residues: 27-217 <NIA>
A;Experimental source: placenta
R;Nir A Bhavnani, N.; Lipner, H.E.
Biochem. Soc. Trans. 1, 205, 1973
A;Title: Catecholamine methyltransferase from human placenta: purification and prop

[illegible]

SEQ
SCH
S:Growth hormone
N:Alternate names: growth hormone
O:Species: Equus caballus (domestic horse)
D:Date: 1-Jul-1981 #sequence_revision 1-Jul-1981 #text_change 23-Aug-1996
A:Accession: A91772; A91495; A91483; A00240; A01514
R:Zukin, M.M.; Poskus, E.; Lanuton, A.A.; Ferrara, P.; Santome, J.A.; Dellacha, J.M.; Pa
Int. J. Pept. Protein Res. 8, 445-444, 1976
A:Title: Primary structure of equine growth hormone.
A:Reference number: A91772; MUID:77005410; PMID:965151
A:Accession: A91772
A:Molecule type: protein
A:Residues: 1-190 <ZAK>
R:Zukin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
[illegible] 54, 49-55, 1973
A:Title: The amino acid sequence of equine growth hormone.
A:Reference number: A91495; MUID:74320362; PMID:4747849
A:Accession: A91495
A:Molecule type: protein
A:Residues: 1-190 <ZAZ>
R:Zukin, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.
Biochem. J. 127, 81-82, 1972
A:Title: Amino acid sequences around the cystine residues in equine growth hormone.
A:Reference number: A91483
A:Accession: A91483
A:Molecule type: protein
A:Residues: 42-69;152-190 <ZA3>
R:Olsen, L.; Hartree, A.S.
Mol. Biol. 3, 144-149, 1968
A:Title: Amino acid sequences around the cystine residues in horse growth hormone.
A:Reference number: A91424 ; MUID:68030830; PMID:5679660
A:Accession: A91424
A:Molecule type: protein
A:Residues: 179-190 <ZLL>
C:Superfamily: prolactin
K:Keywords: hormone; pituitary
F:1-190/Site: SS bond/disulfide bonds: #status experimental

  Query Match 91.8%; Score 159.5; DB 2; Length 190;
  Best Local Similarity 88.1%; Prob. No. 1.5e-42;
  Matches 82; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 180
  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 190
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 190

SEQ
JK:Z19
S:Somatotropin. African elephant
N:Alternate names: growth hormone
O:Species: Loxodonta africana (African elephant)
D:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 15-Dec-1996
A:Accession: JK:Z19
R:Andrews, R.E.; Bloom, M.A.; Li, C.H.; Lam, Y.C.H.
Int. J. Pept. Protein Res. 33, 468-472, 1989
A:Title: Primary structure of elephant growth hormone.
A:Reference number: JK:Z19
A:Accession: JK0219
A:Molecule type: protein
A:Residues: 1-190 <ANL>
A:Experimental source: pituitary gland
C:Superfamily: prolactin
F:1-190/Site: SS bond:somatotropin #status experimental <MA1>

  Query Match 91.8%; Score 159.5; DB 2; Length 190;
  Best Local Similarity 88.1%; Prob. No. 1.5e-42;
  Matches 82; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 180
  FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 190
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQFLRSVFANSLVYGASDSNVYDLLKDLEEGIQTLMGRLEDGSPRTGQIFKQTYSKFDTNSHNDDALLKNYGLLYCFRKDMDKVETFLRIVQCRSVEGSCGF 190

Smith-Waterman [3.1.x]
Copyright (C) 1991 - 2002 Temple F. Smith

Migrated protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 1.4008 Seconds
(without alignments)
4525.0* Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MEVCIPLSKIFNAMLPAAH.....IIIEAYIFKLKVSIL_NE 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41470628 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

1: sp_hum*: 1 of the searched sequences produced by this search have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	256	98.1	217	1	SOMA_HUMAN	P01241 homo sapien
2	255	98.1	217	1	SOMA_MACMU	P28795 macaca mula
3	255	98.1	217	1	SOMA_PANTR	P50754 pan troglod
4	249	95.8	217	1	SOMA_CALJA	Q29HB3 callithrix
5	249	95.8	217	1	SOMA_SAIBO	P50343 saimiri bol
6	239	92.0	217	1	SOMA_PANTR	P50757 pan troglod
7	228	87.7	217	1	SOMA2_HUMAN	P01242 homo sapien
8	199	76.5	217	1	SOMA2_MACMU	Q28779 macaca mula
9	197	75.8	217	1	PRL_HUMAN	P01236 homo sapien
10	163.5	62.9	227	1	SOMA_MOUSE	P06885 mus muscul
11	162.5	62.5	50	1	SOMA_BALMU	P55092 balaenopter
12	161.5	62.1	19	1	SOMA_XENLA	Q28982 xenopus la
13	161.5	62.1	19	1	SOMA_VULVU	P55766 vulpes vulp
14	161.5	62.1	20	1	SOMA_CANFA	P55771 canis famil
15	161.5	62.1	20	1	SOMA_BOSGA	P34834 bos gaurus
16	161.5	62.1	20	1	SOMA_EQUCA	P01215 equus caball
17	161.5	62.1	20	1	SOMA_MESAU	P55767 mesocricetu
18	159.5	61.4	20	1	SOMA_ICTTI	P55249 spermophilu
19	159.5	61.4	216	1	SOMA_RABIT	P46407 oryctolagus
20	159.5	61.4	216	1	SOMA_RAT	P01244 rattus norv
21	159.5	61.4	217	1	SOMA_CAVPO	P16842 cavia porce
22	159.5	61.4	217	1	SOMA_NYCCY	Q9QAB2 nyctereutes
23	159.5	61.4	217	1	SOMA_MUSVI	P19705 mustela vis
24	157.5	60.6	191	1	SOMA_LAMGA	P55765 lama glama
25	156	60.0	216	1	SOMA_MELGA	P22772 meleagris g
26	148	56.9	191	1	SOMA_CHEMY	P49705 chelonia my
27	137	52.7	217	1	SOMA_M_NA	P49706 ondatra zib
28	136	52.3	217	1	SOMA_TRIV	P49707 trichosurus
29	134	51.5	190	1	SOMA_ANTCA	P28773 antidorcas
30	134	51.5	190	1	SOMA_ANTDO	P28774 antidorcas
31	133	51.2	217	1	SOMA_BETOR	P55766 tursiops tr
32	132.5	51.0	217	1	SOMA_BOVIN	P01245 bos taurus
33	132.5	51.0	217	1	SOMA_CEREL	P50437 cervus elap
34	132.5	51.0	217	1	SOMA_CERNI	P50438 cervus nipp
35	132.5	51.0	217	1	SOMA_CERUN	P50439 cervus unic
36	132.5	51.0	217	1	SOMA_ELAHE	P50440 elaphurus d
37	132.5	51.0	217	1	SOMA_GIRCA	P50441 giraffa cam
38	132.5	51.0	217	1	SOMA_HYDHY	P50442 hydropotes
39	132.5	51.0	217	1	SOMA_MOSMO	P50443 moschus mos
40	132.5	51.0	217	1	SOMA_OVIAM	P50444 ovibos mosc
41	132.5	51.0	217	1	SOMA_OVIAR	P50445 ovis aries
42	132.5	51.0	217	1	SOMA_PANHO	P50446 panthalops
43	132.5	51.0	217	1	SOMA_PSEHO	P50447 pseudois no
44	132.5	51.0	217	1	SOMA_RANBE	P50448 rangifer ta
45	132.5	51.0	217	1	SOMA_SYNCA	P50449 syncerus ca

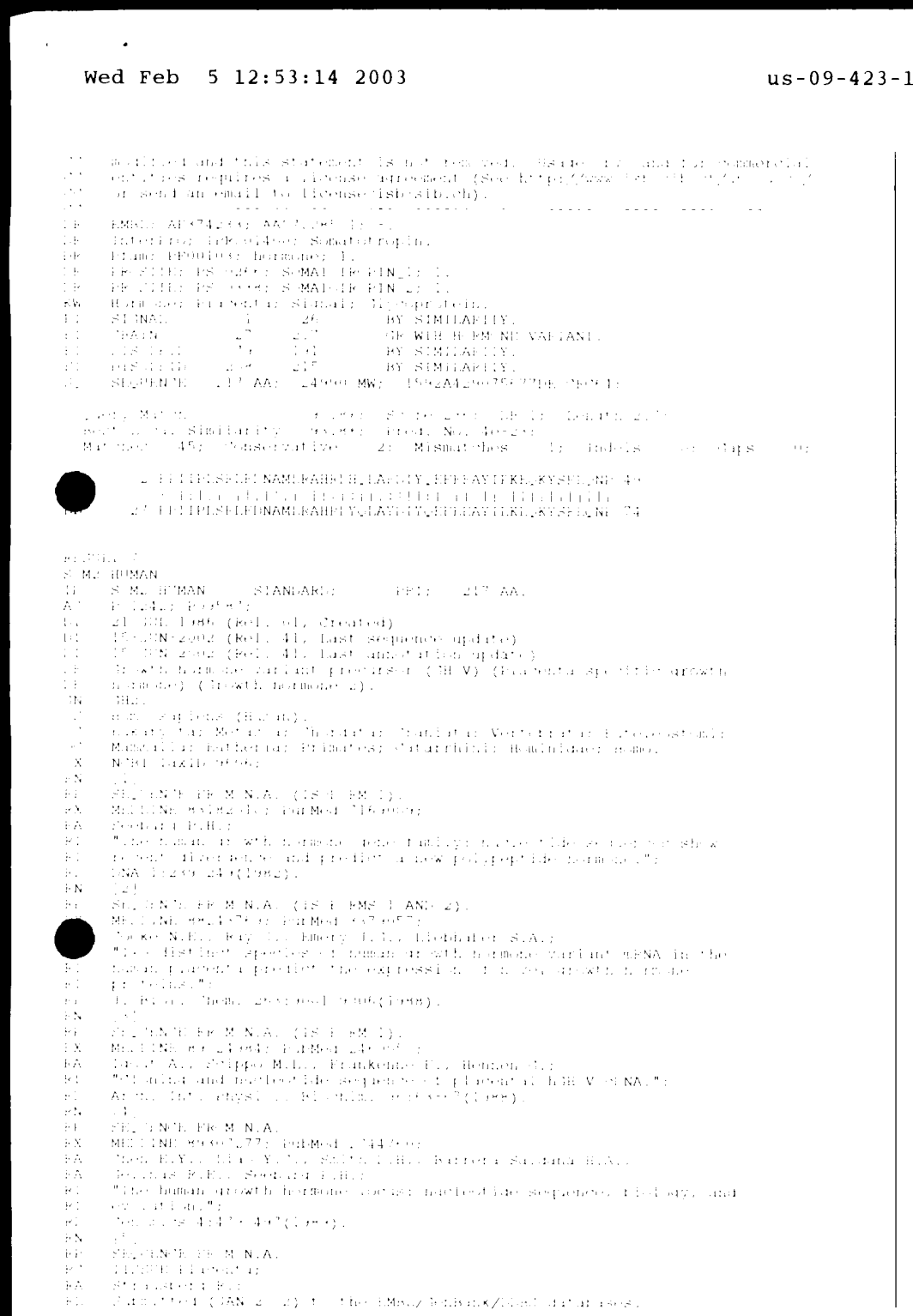
ALIGNMENTS

RESULT 1
SOMA_HUMAN
ID SOMA_HUMAN STANDARD; PRT; 217 AA.
AC P01241; Q13...; Q13...; Q13...; Q13...; Q13...; Q13...;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Somatotropin precursor (Somatotropic hormone) (Growth hormone) (GH).
GN GH1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (CLONE PM...).
RX MEDLINE=86...; PubMed=...;
RA Roskam W., Reuvens H.;
RT "Molecular cloning and characterization of the human growth
RT hormone structural gene.";
RL Nucleic Acids Res. ...(...)...(1986).
RN [2]
RP SEQUENCE FROM N.A. (CLONE PM...).
RX MEDLINE=...; PubMed=...;
RA Martial J.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone: complementary DNA cloning and expression
RT in bacteria.";
RL Science ...(...)...(1979).
RN [3]
RP SEQUENCE FROM N.A. (CLONE PM...), AND POSSIBLE ALTERNATIVE SPLICING.
RX MEDLINE=...; PubMed=...;
RA Seeburg P.H., Martial J.A., Baxter J.D., Goodman H.M.;
RT "Human growth hormone DNA sequence and its expression in cultured
RT mammalian cells.";
RL Nucleic Acids Res. ...(...)...(1981).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA Goodman H.M.;
RT "The human growth hormone gene: structure and evolution.";
RT Recent progress in hormone research ...(...)...(1982).
RL DNA ...(...)...(1982).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA Owerbach D., Lin Y.-S., Shows T.B., Bell G.I., Cohen B.,
RA Goodman H.M.;
RT "The human growth hormone gene: sequence and evolution.";
RT Submitted (SEP-1982) to the EMBL/GenBank/DDBJ databases.
RL

```
RN   [7]
RT   SEQUENCE FROM N.A. (ISOFORM 4).
RA   [illegible]
RX   MEDLINE 20402671; PubMed 10931946;
RA   Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-F., Huang Q.-H., Ren S.-X.,
RA   Gu Y.-J., Huang C.-H., Yu Y.-B., Jiang C.-L., Fu G., Chen Z.-D.,
RA   Ye K.-P., Lu M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA   Xu S.-H., Gu J., Shi J.-X., Jin W., Zhang C.-K., Wu T.-M.,
RA   Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT   "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT   axis and full-length cDNA cloning.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN   [8]
RP   SEQUENCE OF 1-20 FROM N.A.
RX   MEDLINE 86172561; PubMed 3012261;
RA   Gray P.W., Goeddel D.V., McKeown K.S., Derynck R., Chen E.Y.;
RT   "Periplasmic production of correctly processed human growth hormone in
RT   Escherichia coli: natural and bacterial signal sequences are
RT   interchangeable.";
RL   Gene 39:247-254(1985).
RN   [9]
RP   SEQUENCE OF 1-217.
RX   MEDLINE 84208282; PubMed 6418034;
RA   Li C.H., Dixon J.S., Liu W.-K.;
RT   "Human pituitary growth hormone. XIX. The primary structure of the
RT   hormone.";
RL   Arch. Biochem. Biophys. 223:501-508(1983).
RN   [10]
RP   SEQUENCE OF 27-217, AND REVISIONS.
RX   MEDLINE 74145048; PubMed 5143427;
RA   Li C.H., Dixon J.S.;
RT   "Human pituitary growth hormone. XI. The primary structure--a revised
RT   structure.";
RL   Arch. Biochem. Biophys. 148:299-308(1972).
RN   [11]
RP   SEQUENCE OF 27-61 AND 102-134.
RX   MEDLINE 71147565; PubMed 5270048;
RA   Niall H.D.;
RT   "Revised primary structure for human growth hormone.";
RL   Nat. New Biol. 230:90-91(1971).
RN   [12]
RP   REVISIONS TO 119-120 AND 157-159.
RX   MEDLINE 71159048; PubMed 5274528;
RA   Niall H.D., Hogan M.L., Sauer R., Rosenblum I.Y., Greenwood F.C.;
RT   "Sequences of pituitary and placental lactogenic and growth hormones:
RT   evolution from a primordial peptide by gene reduplication.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
RN   [13]
RP   REVISIONS.
RA   Niall H.D.;
RT   "The chemistry of the human lactogenic hormones.";
RT   (In) Griffiths K. (eds.);
RT   Prolactin and carcinogenesis. Proc. fourth Tenovus Workshop, Prolactin,
RL   pp.13-20, Alpha Omega Alpha Press, Cardiff (1972).
RN   [14]
RP   SEQUENCE OF 27-79 (ISOFORM 2).
RX   MEDLINE 81117661; PubMed 7462247;
RA   Chapman G.E., Rogers K.M., Rattigan L., Beardsmore B.A., Bates G.W.,
RA   Lazier C., Cary P.D., Saunders D.C.;
RT   "The 20,000 molecular weight variant of human growth hormone.
RT   Preparation and some physical and chemical properties.";
RL   J. Biol. Chem. 256:8235-8240(1981).
RN   [15]
RP   SEQUENCE OF 16-80 (ISOFORM 2).
RX   MEDLINE [illegible], PubMed [illegible];
RA   [illegible];
RT   "[illegible]";
RL   [illegible].
RN   [16]
RP   REVIEW.
RX   MEDLINE [illegible], PubMed [illegible];
RA   Kaufmann G.;
RT   "Growth hormone heterogeneity in human pituitary and plasma.";
RL   Horm. Res. 51 Suppl. 1:2-7(1999).
RN   [17]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE 90190774; PubMed 1947174;
RA   Cohen F.E., Kuntz I.D.;
RT   "Prediction of the three-dimensional structure of human growth
RT   hormone.";
RL   Proteins 2:162-166(1987).
RN   [18]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX   MEDLINE 92296577; PubMed 1540774;
RA   de Vos A.M., Ultsch M., Kossiakoff A.A.;
RT   "Human growth hormone and extracellular domain of its receptor:
RT   crystal structure of the complex.";
RL   Science 255:306-312(1992).
RN   [19]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE 95375402; PubMed 7684744;
RA   Somers W., Ultsch M., De Vos A.M., Kossiakoff A.A.;
RT   "The X-ray structure of a growth hormone-prolactin receptor complex.";
RL   Nature 372:478-481(1994).
RN   [20]
RP   X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RA   Chantalat L., Chilmaidze N.Y., Dodon N., Korber F., Navaza J.,
RA   Nabel E.R.G., Wlodawer A.;
RT   "The crystal structure of wild-type growth hormone at 2.7 A
RT   resolution.";
RL   Protein Pept. Lett. 2:383-386(1995).
RN   [21]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE 97119042; PubMed 8943043;
RA   Sundstrom M., Lundqvist T., Rosenblad P., Lake M., Edwards B.F.,
RA   Norstedt G.;
RT   "Crystal structure of the antagonist G120R mutant of human growth
RT   hormone, in complex with its receptor at 2.5 A resolution.";
RL   J. Biol. Chem. 271:32797-32803(1996).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of chondrocytes. It also stimulates
CC       cellular uptake of amino acids in muscle and other
CC       tissues.
CC   -!- SUBUNIT: Monomer. Dimer, trimer, tetramer and pentamer. Dimers,
CC       trimers, tetramers, pentamers, hexamers and heptamers and
CC       heteropolymeric combinations. Can also form a complex either with
CC       GHBP or with the alpha-2-acid-glycoprotein complex.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: At least 5 isoforms; 22 kDa (the most
CC       abundant), 20 kDa variant, 6 and 8 are produced by alternative
CC       splicing.
CC   -!- DISEASE: DEFECTS IN GH1 ARE A CAUSE OF PITUITARY DWARFISM I AND
CC       IV.
CC   -!- PHARMACEUTICAL: Available under the names Nutropin (Genentech),
CC       Protropin (Genentech), Norditropin (Novo Nordisk), Genotropin (Pharmacia
CC       Upjohn), Humatrope (Eli Lilly) and Saizen or Serostim (Serono).
CC       Used for treatment of growth hormone deficiency and in
CC       Turner's syndrome.
```

```
ID   [illegible]
AC   [illegible]
DT   [illegible]
DE   Somatotropin precursor (Growth hormone).
GN   GH.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wallis O.C., Wallis M.;
RT   "Characterisation of a primate growth hormone gene";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----------
DR   EMBL; AF159066; AAF01281.1; -.
DR   HSSP; P01241; 1A22.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00337; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    217       SOMATOTROPIN.
FT   DISULFID     79    191       BY SIMILARITY.
FT   DISULFID     81    203       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  24959 MW;  E1D2151A1A7D6142 CRC64;

  Query Match             96.8%;  Score 244;  DB 1;  Length 217;
  Best Local Similarity   97.9%;  Pred. No. 8.7e-25;
  Matches  47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 75


RESULT 2
SOMA_SAIBB
ID   SOMA_SAIBB     STANDARD;      PRT:   217 AA.
AC   P08454;
DT   15-JUN-1988 (Rel. 07, Created)
DT   15-JUN-2002 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH.
OS   Saimiri boliviensis boliviensis (Bolivian squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=39432;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87295486; PubMed=3112982;
RA   Li J.-C., Makova K.D., Adkins R.M., Gibbs S., Li W.-H.;
RT   "Episodic evolution of growth hormone in primates and emergence of the
RT   site-specificity of human growth hormone receptor.";
```

```
RL   Mol. Biol. Evol. 18:1845-1854(2001).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues (By similarity).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----------
DR   EMBL; AF159066; AAF01283.1; -.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00337; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    217       SOMATOTROPIN.
FT   DISULFID     79    191       BY SIMILARITY.
FT   DISULFID     81    203       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  24963 MW;  [illegible] CRC64;

  Query Match             96.8%;  Score 244;  DB 1;  Length 217;
  Best Local Similarity   97.9%;  Pred. No. 8.7e-25;
  Matches  47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 75


RESULT 3
SOMA_PANTR
ID   SOMA_PANTR     STANDARD;      PRT:   217 AA.
AC   P17567;
DT   15-JUN-2000 (Rel. 37, Created)
DT   15-JUN-2002 (Rel. 40, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GH-V) (Placental specific 22 kDa
DE   hormone) (Growth hormone 2).
GN   GH2.
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fox J.A., Espiritu C.L., Sandler L., Porter Saldana B.;
RT   "Independent duplication of the growth hormone gene in three
RT   Anthropoidean lineages.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------- ---- ---- ---- ---- ---- ---- ---- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```



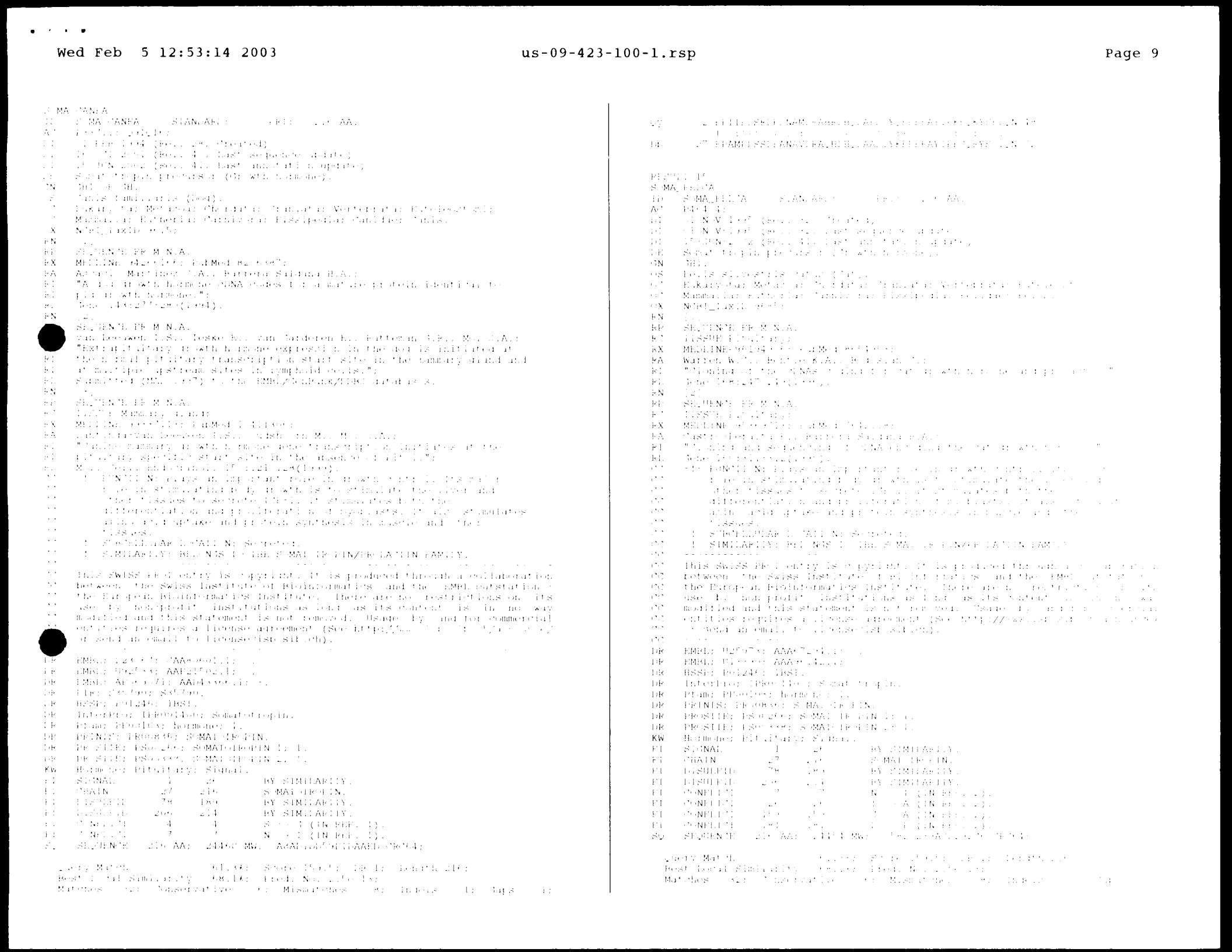
```
AC   xxxxxx; xxxxxx;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GH-V) (Placental specific growth
DE   hormone) (GH with hormone 2).
GN   GH2.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Goles F.G.;
RL   Submitted (JAN 1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=94104617; PubMed=8104617;
RA   Seres E.G., Dutning M., Fisher J.M., Fowler P.B.;
RT   "Cloning of four growth hormone/chorionic somatomammotropin-related
RT   transcriptionary deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey placenta.";
RL   Endocrinology 133:1744-1752(1993).
CC   -!- FUNCTION: Plays an important role in or with control, its major
CC       role in stimulation body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC   -!- SIMILARITY: PHYLOGENETIC TREE AND MULTIPLE ALIGNMENT.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L22464; AAA31131.1; -.
DR   EMBL; L15935; AAA20800.1; -.
DR   HSSI; I01241; IHUMG.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   to STS; hSO2266; SOMATOTROPIN_1; 1.
DR   to STS; hSO4588; SOMATOTROPIN_2; 1.
KW   Hormone; Placenta; Signal; Glycoprotein.
FT   SIGNAL      1     26       BY SIMILARITY.
FT   CHAIN      27    217       GROWTH HORMONE VARIANT.
FT   DISULFID   53    181       BY SIMILARITY.
FT   DISULFID  208    215       BY SIMILARITY.
FT   CONFLICT   57     57       I -> F (IN REF. 2).
FT   CONFLICT  152    152       F -> S (IN REF. 2).
SQ   SEQUENCE   217 AA;  25221 MW;  8C6C1F6B9E24EA00 CRC64;

  Query Match             99.5%;  Score 1090;  DB 2;  Length 217;
  Best Local Similarity   97.1%;  Pred. No. 2.2e-84;
  Matches  211;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MATPLSRIFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
           ||||| |||| |||| |||| |||||||||| | |||||||||||||| |
Db       1 MATPLSRVFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49

RESULT 2
P01241
ID   PRL_HUMAN       STANDARD;       PRT;   217 AA.
AC   P01241;
```

```
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 17, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lactogen precursor (Chorionmammotropin) (Chorionic somatomammotropin) (CS).
GN   CSH1 AND CSH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (GENE CSH1).
RX   MEDLINE=85036126; PubMed=6208792;
RA   Selby M.J., Barta A., Boorstein W.R., Pelham R.J., Eberhardt N.L.;
RT   "Analysis of a major human chorionic somatomammotropin gene: evidence
RT   for two functional promoter elements.";
RL   J. Biol. Chem. 259:4405-4412(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (GENE CSH1).
RX   MEDLINE=87161235; PubMed=3549690;
RA   Hill R., Kimelman D., Birnbaum M.J., Baxter J.D., Seeburg P.H.,
RA   Eberhardt N.L., Barta A.;
RT   "The human growth hormone gene locus: structure, evolution, and
RT   allelic variations.";
RL   DNA 6:59-70(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86160016; PubMed=3456456;
RA   Barrera Saldana H.A., Seeburg P.H., Saunders G.F.;
RT   "Two structurally different genes produce the same secreted human
RT   placental lactogen hormone.";
RL   J. Biol. Chem. 258:3787-3793(1983).
RN   [4]
RP   SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).
RX   MEDLINE=...2477; PubMed=...;
RA   Chen E.Y., Liao Y.C., Smith D.H., Barrera Saldana H.A., Gelinas R.E.,
RA   Seeburg P.H.;
RT   "The human growth hormone ... sequence: ... sequence, blood ... and
RT   evolution.";
RL   Genomics 4:479-497(1989).
RN   [5]
RP   SEQUENCE.
RX   MEDLINE=84182010; PubMed=...;
RA   Seeburg P.H.;
RT   "The human growth ... family: ... Alu sequences and
RT   recent divergence and predict a new polypeptide hormone.";
RL   DNA 1:239-249(1982).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Selby M.J.;
RL   Submitted (JAN 1990) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 5-217 FROM N.A.
RX   MEDLINE=...; PubMed=...;
RA   Shine J., Seeburg P.H., Martial J.A., Baxter J.D., Goodman H.M.;
RT   "Construction and analysis of recombinant DNA for human ...
RT   somatomammotropin.";
RL   Nature 270:494-499(...).
RN   [8]
RP   SEQUENCE OF 27-217.
RX   MEDLINE=...; PubMed=...;
RA   Li C.H., Dixon J.S., Chung D.;
RT   "Amino acid sequence of human chorionic somatomammotropin.";
RL   Arch. Biochem. Biophys. 1:...-... (...).
RN   [9]
RP   SEQUENCE OF 27-217.
RX   MEDLINE=...; PubMed=...;
RA   Shewale J.G., Baldwin S., McCarthy W., ..., Talmadge K.;
RT   "Amino acid sequence of human placental lactogen.";
RL   Nature New Biol. ...:...-... (...).
RN   [10]
RP   ERRATUM.
RA   Shewale J.G., Baldwin S., McCarthy W., Talmadge K.;
```

```
ID   Native New Biol. 235:64-64(1972).
XX   ;;;
RA   "Identification of the interchain disulfide bonds of dimeric human
RT   placental lactogen.";
RL   ;;; 2:54, 2002 (1972).
XX     1  INCLUDE: Swiss-Prot is copyright. It is produced through a collab
XX     2  between the Swiss Institute of Bioinformatics and the EMBL outstation -
XX     3  the European Bioinformatics Institute. There are no restrictions on its
XX     4  use by non-profit institutions as long as its content is in no way
XX     5  modified and this statement is not removed. Usage by and for commercial
XX     6  entities requires a license agreement (See http://www.isb-sib.ch/announce/
XX     7  or send an email to license@isb-sib.ch).
XX   ---------- ---------- ---------- ---------- ---------- ---------- ----------
XX
XX   EMBL: V00570; CAA23836.1;  -
XX   EMBL: J02871; AAA40747.1;  -
XX   EMBL: K02434; AAA52139.1;  -
XX   EMBL: M57441; AAA52139.1;  -
XX   EMBL: X00076; AAA52551.1;  -
XX   EMBL: J00119; AAA98421.1;  -
XX   EMBL: BC005021; AAH05021.1;  -
XX   EMBL: BC020756; AAH20756.1;  -
XX   PIR: A25122; GHBP.
XX   PIR: A26440; A26440.
XX   PIR: C32435; C32435.
XX   PIR: E32435; E32435.
XX   HSSP: P01241; 1A22
XX   InterPro: IPR000441; SOMATOTROPIN.
XX   Pfam: PF00103; hormone_1;
XX   PRINTS: PR00837; SOMATOTROPIN.
XX   PROSITE: PS00264; SOMATOTROPIN_1; 1.
XX   PROSITE: PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Multigene family; Signal.
FT   SIGNAL       1     26
FT   CHAIN       27    217       GROWTH HORMONE.
FT   DISULFID    53    165
FT   DISULFID   208    215
FT   DISULFID   208    208       INTERCHAIN (WITH P42 IN A DIMER).
FT   DISULFID   215    215       INTERCHAIN (WITH P42 IN A DIMER).
FT   VARIANT      1      1       A -> T (IN PSH-).
FT                               /FTId=VAR_007167.
FT   VARIANT    164    164       IS -> L (IN PSH-).
FT                               /FTId=VAR_007168.
FT   CONFLICT    91     94       I -> T (IN REF. 9).
FT   CONFLICT    95     95       MISSING (IN REF. 7).
FT   CONFLICT   119    119       MISSING (IN REF. 7).
FT   CONFLICT   144    145       SD -> RS (IN REF. 9).
SQ   SEQUENCE   217 AA;  25028 MW;  2458D0D7A7184E41 CRC64;

  Query Match             95.4%;  Score 1675;  DB 1;  Length 217;
  Best Local Similarity   99.5%;  Pred. No. 1.1e-164;
  Matches  216;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
```

```
ID   ol_JAN-1998 (Rel. 4, Created)
DT   15-JUN-2000 (Rel. 38, Last sequence update)
DT   15-JUN-2000 (Rel. 38, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sinha A.M., Vasavada R.C., Asch B.B.;
RT   "Structure of the mouse growth hormone gene and its pituitary
RT   mRNA.";
RL   Submitted (JUL-1988) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Growth hormone is important for growth, cell
CC       differentiation and pituitary hormone regulation.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   ---------- ---------- ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ---------- ---------- ----------
DR   EMBL; X00497; CAA25136.1;  -
DR   EMBL; Z46667; CAA86477.1;  -
DR   PIR; B23911; SOMS.
DR   HSSP; P01241; 1A22.
DR   MGD; MGI:97164; Gh.
DR   InterPro; IPR000441; SOMATOTROPIN.
DR   Pfam; PF00103; hormone_1;
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26       BY SIMILARITY.
FT   CHAIN       27    217       GROWTH HORMONE.
FT   DISULFID    53    165       BY SIMILARITY.
FT   DISULFID   208    215       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  25028 MW;  2458D0D7A7184E41 CRC64;

  Query Match             94.5%;  Score 1659;  DB 1;  Length 217;
  Best Local Similarity   98.6%;  Pred. No. 3.1e-163;
  Matches  214;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP
```

```
DE   Somatotropin (Growth hormone).
GN   Sat.
OS   Balaenoptera borealis (Sei whale).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC   Balaenopteridae; Balaenoptera.
OX   NCBI_TaxID=9768;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=85*>; PubMed=2115*1;
RA   Tedesco N.A., Ponkov I.A., Bolatov A.A., Stipova I.A.;
RT   "Amino acid sequence of seiwhale somatotropin.";
RL   Biokhimiia 4*:1059-1*69(1985).
RN   [2]
RP   PRELIMINARY PARTIAL SEQUENCE.
RA   Stipova I.A., Bolatov A.A., Ponkov Y.A.;
RL   "Structural studies of tryptic peptides from cyanogen bromide
RL   treatments of sei whale (Balaenoptera borealis) somatotropin.";
RL   Biokhim. Khim. 4:1585 (1978).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; JN0440; JN0440.
DR   PIR; JN0487; JN0487.
DR   HSSP; P01241; 1AX1.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; Hormone_1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     53     183       BY SIMILARITY.
FT   DISULFID     189    199       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21761 MW;  9DFBFF0DB11A7560 CRC64;

  Query Match             61.7%;  Score 760.5;  DB 1;  Length 190;
  Best Local Similarity   98.1%;  Pred. No. 1.6e-1;
  Matches  182;  Conservative    5;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQN 48
           ||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db       1 FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPKDQRYTFLQN 48


RESULT 13
SOMA_VULVU
ID   SOMA_VULVU     STANDARD;      PRT;   190 AA.
AC   P01244;
DT   01-JUL-1987 (Rel. 11, Created)
DT   01-JUL-1987 (Rel. 11, Last sequence update)
DT   15-JUN-2004 (Rel. 43, Last annotation update)
DE   Somatotropin (Growth hormone).
GN   GH1.
OS   Vulpes vulpes (Red fox).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Vulpes.
OX   NCBI_TaxID=9627;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pituitary;
RX   MEDLINE=89254275; PubMed=2722401;
RA   Li C.H., Dzhobski J., Chung D.;
RT   "Primary structure of fox pituitary growth hormone.";
RL   Int. J. Pept. Protein Res. 33:78-82(1989).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; X54120; X54120.
DR   HSSP; P01241; 1BSI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; Hormone_1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     53     183       BY SIMILARITY.
FT   DISULFID     189    199       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21761 MW;  9DFBFF0DB11A7560 CRC64;

  Query Match             61.7%;  Score 760.5;  DB 1;  Length 190;
  Best Local Similarity   98.1%;  Pred. No. 1.6e-1;
  Matches  182;  Conservative    5;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQN 48
           ||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db       1 FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPKDQRYTFLQN 48


RESULT 14
```



Title: US-09-423-100-1
Perfect score: 260
Sequence: 1 MSTTPPLSPLPTNAMKRAHE.....IPTLAYTPE_KYSTL_NP 49

Searched: 2750850 seqs, 2 x 847115 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_rodent:*
8: sp_virus:*
9: sp_vertebrate:*
10: sp_unclassified:*
11: sp_rnavirus:*
12: sp_typesequences:*
13: sp_patentseqs:*

Total No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	260	100.0	47	2WKN_	2WKN_ Protein acet
2	260	100.0	47	1	Q4Y4I3 Putative
3	260	100.0	212	1	Q97XY9 Haloarcula
4	260	100.0	212	1	Q97XY8 Haloarcula
5	260	100.0	212	1	Q97XY7 Haloarcula
6	260	100.0	212	1	Q97XY6 Haloarcula
7	260	100.0	212	1	Q97XY5 Haloarcula
8	260	100.0	212	1	Q97XY4 Haloarcula
9	260	100.0	212	1	Q97XY3 Haloarcula
10	260	100.0	212	1	Q97XY2 Haloarcula
11	260	100.0	212	1	Q97XY1 Haloarcula
12	260	100.0	212	1	Q97XY0 Haloarcula
13	260	100.0	212	1	Q97XX9 Haloarcula
14	260	100.0	212	1	Q97XX8 Haloarcula
15	260	100.0	212	1	Q97XX7 Haloarcula
16	260	100.0	212	1	Q97XX6 Haloarcula
17	260	100.0	212	1	Q97XX5 Haloarcula
18	260	100.0	212	1	Q97XX4 Haloarcula
19	260	100.0	212	1	Q97XX3 Haloarcula
20	260	100.0	212	1	Q97XX2 Haloarcula
21	260	100.0	212	1	Q97XX1 Haloarcula
22	260	100.0	212	1	Q97XX0 Haloarcula
23	260	100.0	212	1	Q97XW9 Haloarcula
24	260	100.0	212	1	Q97XW8 Haloarcula
25	260	100.0	212	1	Q97XW7 Haloarcula
26	260	100.0	212	1	Q97XW6 Haloarcula
27	260	100.0	212	1	Q97XW5 Haloarcula
28	260	100.0	212	1	Q97XW4 Haloarcula
29	260	100.0	212	1	Q97XW3 Haloarcula
30	260	100.0	212	1	Q97XW2 Haloarcula
31	260	100.0	212	1	Q97XW1 Haloarcula
32	260	100.0	212	1	Q97XW0 Haloarcula
33	260	100.0	212	1	Q97XV9 Haloarcula
34	260	100.0	212	1	Q97XV8 Haloarcula
35	260	100.0	212	1	Q97XV7 Haloarcula
36	260	100.0	212	1	Q97XV6 Haloarcula
37	260	100.0	212	1	Q97XV5 Haloarcula
38	260	100.0	212	1	Q97XV4 Haloarcula
39	260	100.0	212	1	Q97XV3 Haloarcula
40	260	100.0	212	1	Q97XV2 Haloarcula
41	260	100.0	212	1	Q97XV1 Haloarcula
42	260	100.0	212	1	Q97XV0 Haloarcula
43	260	100.0	212	1	Q97XU9 Haloarcula
44	260	100.0	212	1	Q97XU8 Haloarcula
45	260	100.0	212	1	Q97XU7 Haloarcula

ALIGNMENTS

RESULT 1
2WKN_A
ID 2WKN_A STANDARD; PRT; 47 AA.
AC 2WKN_A;
DT 01-MAR-2005 (TREMBLrel. 28, Created)
DT 01-MAR-2005 (TREMBLrel. 28, Last sequence update)
DT 01-JUN-2005 (TREMBLrel. 30, Last annotation update)
DE Putative transposase.
GN ORF1.
OS Alcelaphine herpesvirus 1 (Malignant catarrhal fever virus).
OC Viruses; Herpesvirales; Herpesviridae; Gammaherpesvirinae;
OC Macavirus.
OX NCBI_TaxID=10385;
RN [1]
RP SEQUENCE FROM N.A.
RA Bowden A.R., Estes D.M., Schultz E.R., Barrett T., Reid H.W.,
RA Li H.;
RT "Independent segregation of two strains of alcelaphine herpesvirus 1
RT Athraxpoidoum tincorpos.";
RL Submitted (APR-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF377...; AAL...1.1; -.
DR InterPro; IPR...; Schaf...-like.
DR Pfam; PF...; Schaf...; 1.
DR PRINTS; PR...; Schaf...FAMILY.
DR PROSITE; PS...; Schaf...1, Schaf...2, Schaf...3.
DR PROSITE; PS...; Schaf...1, Schaf...2, Schaf...3.
SQ SEQUENCE 47 AA; 4...6 MW; 4...6...AAAA...6 CRC64;

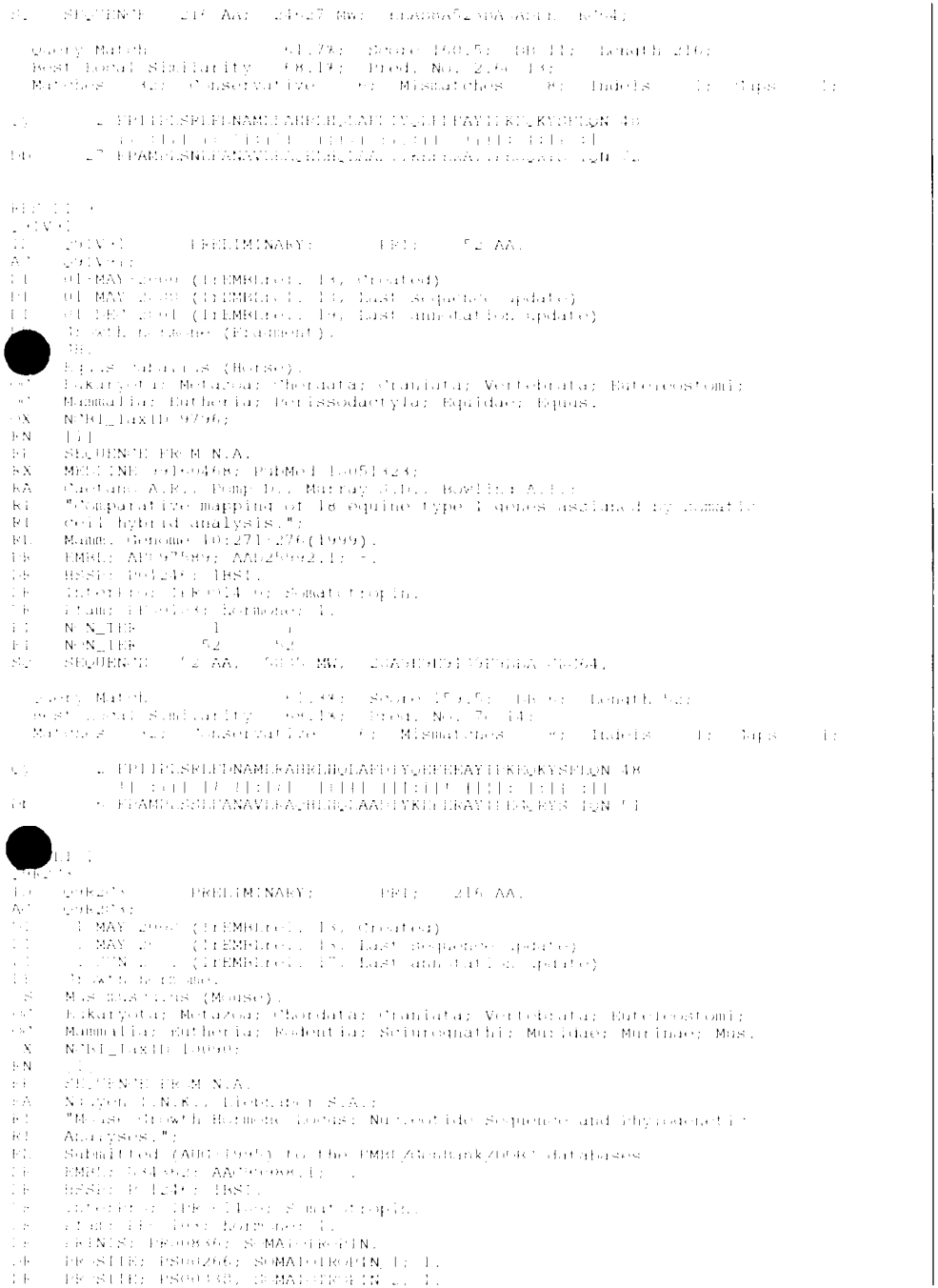
Query Match ... Score ...; ... Length ...; ...
Best Local Similarity ...; ... Conservative ...; ... Mismatches ...; ... Indels ...; ... Gaps ...;
Matches ...; ... Conservative ...; ... Mismatches ...; ... Indels ...; ... Gaps ...;

Qy ... MSTTPPLSPLPTNAMKRAHE...IPTLAYTPE_KYSTL_NP 49

Db ... MSTTPPLSPLPTNAMKRAHE...IPTLAYTPE_KYSTL_NP 47

RESULT 2
2WKN_B
ID 2WKN_B STANDARD; PRT; 47 AA.

```
AC   I46411;
DT   01-JAN-1988 (IrEMBLrel. 05, Created)
DT   01-JAN-1988 (IrEMBLrel. 05, Last sequence update)
DT   01-OCT-2001 (IrEMBLrel. 19, Last annotation update)
DE   Placental growth hormone isoform hPGH-V precursor.
GN   HGH-V.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=91008721; PubMed=2404017;
RA   Owerbach D., Bell G.I., Rutter W.J., Shows T.B.;
RT   "Clustering of four growth hormone/chorionic somatomammotropin related
RT   complementary deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey placenta.";
RL   Endocrinology 128:1744-1752(1991).
DR   EMBL; M59958; AAA31839.1; -.
DR   PIR; A25744; GHHUV5.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
KW   Pituitary; Hormone; Signal.
FT   SIGNAL       1      26       POTENTIAL.
SQ   SEQUENCE   245 AA;  27191 MW;  1F7F7F9C0191B7B8 CRC64;

  Query Match             87.7%;  Score 228;  DB 1;  Length 245;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-22;
  Matches  44;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
           ||||||||||||||| ||||||||||||| |||||||||||
Db      27 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 76


RESULT 4
Q07367
ID   Q07367      PRELIMINARY;      PRT;   217 AA.
AC   Q07367;
DT   01-NOV-1996 (IrEMBLrel. 04, Created)
DT   01-NOV-1996 (IrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE   Somatotropin 1 precursor (Growth hormone 1).
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=91008721; PubMed=2404017;
RA   Owerbach D., Bell G.I., Rutter W.J., Shows T.B.;
RT   "Cloning of four growth hormone/chorionic somatomammotropin related
RT   complementary deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey placenta.";
RL   Endocrinology 128:1744-1752(1991).
CC   -!- FUNCTION: STIMULATES BONE AND CARTILAGE GROWTH.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
DR   EMBL; L16752; AAA31839.1; -.
DR   BSSP; PS1241; IAX1.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PROSITE; PS00266; SOMATOTROPIN_2; UNKNOWN 1.
KW   Pituitary; Hormone; Signal.
FT   SIGNAL       1      2
FT   CHAIN        3    217       SOMATOTROPIN 1.
FT   DISULFID     71    181       BY SIMILARITY.
FT   DISULFID    208    215       BY SIMILARITY.
SQ   SEQUENCE   217 AA;  23912 MW;  1F1AA4C2C1C1E2B6 CRC64;

  Query Match             81.9%;  Score 214;  DB 6;  Length 217;
  Best Local Similarity   78.7%;  Pred. No. 2.6e-20;
  Matches  48;  Conservative    9;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 PTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
           |||||||||||||| |||||||||||||||||||||||||||||||||
Db      28 PTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 76


RESULT 5
Q07369
ID   Q07369      PRELIMINARY;      PRT;   217 AA.
AC   Q07369;
DT   01-NOV-1996 (IrEMBLrel. 04, Created)
DT   01-NOV-1996 (IrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE   Somatotropin 3 precursor (Growth hormone 3).
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=94009724; PubMed=8144617;
RA   Shows T.B., Daniels M., Fisher J.M., Rutter W.J.;
RT   "Cloning of four growth hormone/chorionic somatomammotropin related
RT   complementary deoxyribonucleic acids differentially expressed during
RT   pregnancy in the rhesus monkey placenta.";
```



```
ID   N_N_IEP     17%    17%
SQ   SEQUENCE   17% AA;  2xxxx MW;  [illegible]

  Query Match             58.5%;  Score 152;  DB 6;  Length 17%;
  Best Local Similarity   65.1%;  Pred. No. 2.6e-12;
  Matches   28;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYS 44
           |:|:|||:||  |||||  |:|||  |||:||  :|||:::  :
Db       6 FDSVPLSKLFANAVLRAQHLHQLAADTYKEFERIYIPEGQRNS 50


RESULT 14
Q9SMJ6
ID   Q9SMJ6      PRELIMINARY;      PRT;   204 AA.
AC   Q9SMJ6;
DT   01-DEC-2000 (TrEMBLrel. 16, Created)
DT   01-DEC-2000 (TrEMBLrel. 16, Last sequence update)
DT   09-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Growth hormone (Fragment).
OS   Tursiops aduncus (Indian).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Tursiops.
OX   NCBI_TaxID=9739;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21205545; PubMed=11567041;
RA   Liu J.D., Makova K.D., Adkins R.M., Gibson S., Li W.H.;
RT   "Episodic Evolution of Growth Hormone in Primates and Emergence of the
RT   Species Specificity of Human Growth Hormone Receptor.";
RL   Mol. Biol. Evol. 18:2087-2094(2001).
DR   EMBL; AF339066; AAK62307.1; -.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00340; SOMATOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   204 AA;  23492 MW;  [illegible]

  Query Match             57.5%;  Score 149.5;  DB 6;  Length 204;
  Best Local Similarity   63.8%;  Pred. No. 7e-12;
  Matches   30;  Conservative    7;  Mismatches    9;  Indels    1;  Gaps    1;

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQN 48
           |:|:|||:||  |||||  ||||:  |||:||  ||:|  |:|:|||
Db       6 FDAMPLSKLFANAVLRAQHLMQLAADTFKEFERTYIPEGQRTSIQN 50


RESULT 15
  ...
ID   ...:      PRELIMINARY;      PRT;   202 AA.
AC   ...;
DT   ...-JAN-1998 (TrEMBLrel. 05, Created)
DT   ...-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   ...-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Placental growth hormone 20Kda isoform precursor.
GN   HGH-V.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=99373737; PubMed=9927665;
RA   [illegible]
RA   [illegible]
RT   "Isolation of two novel growth hormone transcripts expressed in human
RT   placenta.";
RL   J. Clin. Endocrinol. Metab. [illegible]
DR   EMBL; [illegible]
DR   EMBL; AF[illegible]
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00340; SOMATOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   202 AA;  [illegible]

  Query Match             [illegible]  Score 149;  DB 6;  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches   28;  Conservative    7;  Mismatches    9;  Indels    1;  Gaps    1;

QY       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYS 44
           |:|:|||:||  |||||  |:|||  |||:||  :|||:::  :
Db       6 FDSVPLSKLFANAVLRAQHLHQLAADTYKEFERIYIPEGQRNS 50


Search completed: February [illegible], 2003, [illegible]
Job time : [illegible] secs
```

MPsearch protein search, using sw model

Run on: February 5, 2003, 12:44:51 ; Search time 1 Seconds
(without alignments)
24.459 Million cell updates/sec

Title: US-09-423-100-1
Perfect score: 2[illegible]
Sequence: 1 MPSS[illegible]NAMQIAAN[illegible]...[illegible]QIRESSYIER_ESSQIL[illegible] 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: [illegible] seqs, [illegible] residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubaa/US06_PUB.pep:*
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubaa/US07_PUB.pep:*
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubaa/US08_PUB.pep:*
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubaa/US09_PUB.pep:*
9: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubaa/US10_PUB.pep:*
11: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubaa/US11_PUB.pep:*
13: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubaa/US12_PUB.pep:*
15: /cgn2_6/ptodata/1/pubaa/US12_NEW_PUB.pep:*

Total No. is the number of results probably by chance
with the greater than or equal to the score of the hit as being printed,
and is obtained by analysis of the total score distribution.

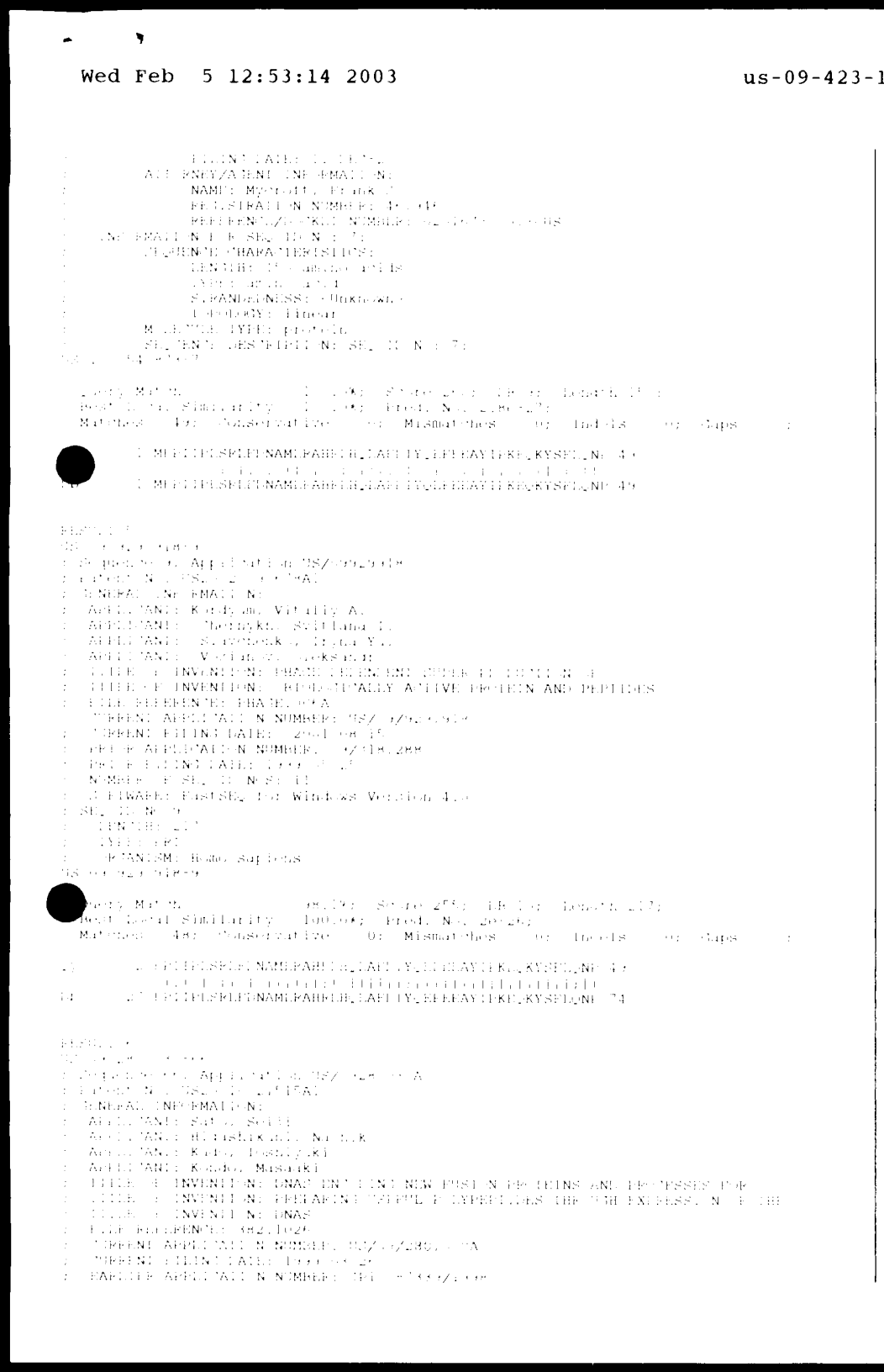
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 2, Appli
3	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 5, Appli
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 7, Appli
5	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 6, Appli
7	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 2, Appli
8	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 10, Appl
9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 4, Appli
10	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 3, Appli
11	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 14, Appl
12	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 2, Appli
13	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 4, Appli
14	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 7, Appli
15	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 6, Appli
16	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 2, Appli
17	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 12, App
18	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
19	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
20	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
21	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
22	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
23	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
24	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
25	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
26	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
27	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
29	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
30	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
31	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
32	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
33	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
34	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
35	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
36	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
37	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
38	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
39	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
40	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
41	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
42	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
44	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli
45	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-423-100-1
; Sequence 1, Application US/09423100
; Patent No. US6228611
; GENERAL INFORMATION:
; APPLICANT: [illegible]
; TITLE OF INVENTION: [illegible]
; NUMBER OF SEQUENCES: [illegible]
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: [illegible]
; CITY: San Francisco
; STATE: [illegible]
; COUNTRY: USA
; ZIP: 94111-4065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: [illegible] Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: [illegible]
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: [illegible]
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible] amino acids
; TYPE: amino acid
; STRANDEDNESS: [illegible]
; TOPOLOGY: [illegible]
; MOLECULE TYPE: [illegible]
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
RESULT 1
US-09-423-100-1
; Sequence 1, Application US/09423100
; Patent No. US6054287
;  GENERAL INFORMATION:
;       APPLICANT: Gan, Zhong Hu
;       TITLE OF INVENTION: Chimeric Protein Containing an
;                           Intramolecular Chaperone Like Sequence
;       NUMBER OF SEQUENCES: 7
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/423,100
;            FILING DATE: 22-Jan-2002
;            CLASSIFICATION: Unknown
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/871,827
;            FILING DATE: 09-JUN-1997
;            APPLICATION NUMBER: US 08/423,100
;            FILING DATE: 11-DEC-2000
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mycroft, Frank J
;            REGISTRATION NUMBER: 40,946
;            REFERENCE/DOCKET NUMBER: 020167-000100US
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 92 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: Unknown
;            TOPOLOGY: Linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-423-100-1

  Query Match             100.0%;  Score 460;  DB 4;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-27;
  Matches  460;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MEPETIPLSKLFNAMLRARRLQLAFQIYGEEGAVTPKFLKYSPLQNFI 47
            |||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MEPETIPLSKLFNAMLRARRLQLAFQIYGEEGAVTPKFLKYSPLQNFI 47


RESULT 2
US-09-423-100-2
; Sequence 2, Application US/09423100
; Patent No. US6054287
;  GENERAL INFORMATION:
;       APPLICANT: Gan, Zhong Hu
;       TITLE OF INVENTION: Chimeric Protein Containing an
;                           Intramolecular Chaperone Like Sequence
;       NUMBER OF SEQUENCES: 7
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/423,100
;            FILING DATE: 22-Jan-2002
;            CLASSIFICATION: Unknown
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/871,827
;            FILING DATE: 09-JUN-1997
;            APPLICATION NUMBER: US 08/423,100
;            FILING DATE: 11-DEC-2000
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Mycroft, Frank J
;            REGISTRATION NUMBER: 40,946
;            REFERENCE/DOCKET NUMBER: 020167-000100US
;   INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 92 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: Unknown
;            TOPOLOGY: Linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-423-100-2

  Query Match             100.0%;  Score 460;  DB 4;  Length 460;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-27;
  Matches  460;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 MEPETIPLSKLFNAMLRARRLQLAFQIYGEEGAVTPKFLKYSPLQNFI 47
            |||||||||||||||||||||||||||||||||||||||||||||||
Db        1 MEPETIPLSKLFNAMLRARRLQLAFQIYGEEGAVTPKFLKYSPLQNFI 47


RESULT 3
US-09-423-100-3
; Sequence 3, Application US/09423100
; Patent No. US6054287
;  GENERAL INFORMATION:
;       APPLICANT: Gan, Zhong Hu
;       TITLE OF INVENTION: Chimeric Protein Containing an
;                           Intramolecular Chaperone Like Sequence
;       NUMBER OF SEQUENCES: 7
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Townsend and Townsend and Crew LLP
;            STREET: Two Embarcadero Center, Eighth Floor
;            CITY: San Francisco
;            STATE: California
;            COUNTRY: USA
;            ZIP: 94111-3834
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: Patentin Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/423,100
;            FILING DATE: 22-Jan-2002
;            CLASSIFICATION: Unknown
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/871,827
;            FILING DATE: 09-JUN-1997
;            APPLICATION NUMBER: US 08/423,100
;            FILING DATE: 11-DEC-2000
```



```
  Best Local Similarity   97.9%;  Pred. No. 1.7e-25;
  Matches   47;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db         27 FPTIPLSRPFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74


RESULT 9
US-09-853-688-4
; Sequence 4, Application US/09853688
; Patent No. US20020081605A1
; GENERAL INFORMATION:
;  APPLICANT: COOPER, DAVID N.
;  APPLICANT: PROCTER, ANNIE M.
;  APPLICANT: GREGORY, JOHN
;  APPLICANT: MILLAR, DAVID S.
;  TITLE OF INVENTION: METHOD FOR DETECTING GROWTH HORMONE VARIATIONS IN
;  TITLE OF INVENTION: HUMANS, THE VARIATIONS AND THEIR USES
;  FILE REFERENCE: WCM78
;  CURRENT APPLICATION NUMBER: US/[illegible]
;  CURRENT FILING DATE: 2001-0[illegible]-14
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 217
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-853-688-4

  Query Match             93.1%;  Score 242;  DB 10;  Length 217;
  Best Local Similarity   95.8%;  Pred. No. 1e-24;
  Matches   46;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db         27 FPTIPLSRLFDNASLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 74


RESULT 10
US-09-850-887-3
; Sequence 3, Application US/09850887
; Patent No. US20020[illegible]778A1
;    GENERAL INFORMATION:
;         APPLICANT: Lal, Preeti
;                    Corley, Neil C.
;                    Tang, Y. Tom
;         TITLE OF INVENTION: THYROID AND PITUITARY HORMONE PROTEIN
;         NUMBER OF SEQUENCES: 4
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
;              STREET: 3174 PORTER DRIVE
;              CITY: PALO ALTO
;              STATE: CALIFORNIA
;              COUNTRY: USA
;              ZIP: 94304
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: IBM PC compatible
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/[illegible]
;              FILING DATE: 07-May-2001
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: [illegible]
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: FERRONE, MICHAEL C.
;              REGISTRATION NUMBER: [illegible]
;              REFERENCE/DOCKET NUMBER: PF-[illegible] US
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (650) 855-0555
;              TELEFAX: (650) 845-4166
;    INFORMATION FOR SEQ ID NO: 3:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 217 amino acids
;              TYPE: amino acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         IMMEDIATE SOURCE:
;              LIBRARY: GenBank
;              CLONE: g4[illegible]
;         SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-850-887-3

  Query Match             76.[illegible]%;  Score [illegible];  DB 10;  Length 217;
  Best Local Similarity   7[illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNP 49
Db         27 [illegible]KYSFLRNP 74


RESULT 11
US-09-876-478-14
; Sequence 14, Application US/09876478
; Patent No. US2002[illegible]214A1
; GENERAL INFORMATION:
;  APPLICANT: Peyman, John A.
;  TITLE OF INVENTION: Interferon Suppressing Placental Lactogen Peptides
;  FILE REFERENCE: 1161
;  CURRENT APPLICATION NUMBER: US/09/876,478
;  CURRENT FILING DATE: [illegible]
;  PRIOR APPLICATION NUMBER: 60/210,[illegible]
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: MS-DOS
; SEQ ID NO 14
;   LENGTH: 54
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
;   FEATURE:
;   OTHER INFORMATION: N TERMINAL 54 RESIDUES OF [illegible]
US-09-876-478-14

  Query Match             54.6%;  Score 1[illegible];  DB [illegible];  Length 54;
  Best Local Similarity   9[illegible]%;  Pred. No. [illegible];
  Matches   27;  Conservative    0;  Mismatches    1;  Indels    [illegible];  Gaps    [illegible];

Qy          2 FPTIPLSRLFDNAMLRAHRLHQLAFDTY 29
Db         27 FPTIPLSRLFDNASLRAHRLHQLAFDTY 54


RESULT 12
US-09-876-478-2
; Sequence 2, Application US/09876478
; Patent No. US20020[illegible]214A1
; GENERAL INFORMATION:
;  APPLICANT: Peyman, John A.
;  TITLE OF INVENTION: Interferon Suppressing Placental Lactogen Peptides
;  FILE REFERENCE: 1161
;  CURRENT APPLICATION NUMBER: US/09/876,478
;  CURRENT FILING DATE: 2001-[illegible]
;  PRIOR APPLICATION NUMBER: 60/210,082
;  PRIOR FILING DATE: 2000-06-07
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: MS-DOS
; SEQ ID NO 2
;   LENGTH: 54
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
;   FEATURE:
```

[illegible]

Run on: February 5, 2003, 12:42:21 ; Search time 7.5911 seconds
(without alignments)
52.115 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MTEITIGSRLPTNAMLRAIR.....NGEIIRISIIISWSIIVL 92

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 59470 seqs, 16250620 residues

Total number of hits satisfying chosen parameters: 59470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_main.seq:*
2: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_new.seq:*
3: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_upd.seq:*
4: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_var.seq:*
5: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_wo.seq:*
6: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_wo_new.seq:*
7: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_wo_upd.seq:*
8: /SIDS2/gcgdata/geneseq/geneseq-embl/AA_wo_var.seq:*

Total No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is estimated by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	470	100.0	92	2	AAY42267	Human growth hormo
2	470	100.0	184	2	AAW42267	Human anti-androge
3	470	100.0	190	2	AAY42268	Chimeric protein
4	468	99.6	160	1	AAP01941	Human GH with heter
5	468	99.6	162	1	AAP01942	Human GH with heter
6	468	99.6	176	2	AAW42268	Human anti-androge
7	468	99.6	191	2	AAY42269	Human GH with heter
8	468	99.6	262	2	AAB71781	Human GH with heter
9	468	99.6	320	11	AAE42525	First G protein-li
10	466	99.2	144	11	AAP05615	Segment of Bovell
11	464	98.7	92	1	AAP01940	Human GH with heter
12	464	98.7	92	2	AAW42266	Human anti-androge
13	464	98.7	92	2	AAY42266	Human growth hormo
14	464	98.7	92	2	AAY42270	Human growth hormo
15	464	98.7	92	2	AAY42271	Human growth hormo
16	464	98.7	92	2	AAY42272	Human growth hormo
17	464	98.7	92	2	AAY42273	Human growth hormo
18	464	98.7	184	2	AAW42268	Human anti-androge
19	464	98.7	184	2	AAW42269	Human anti-androge
20	464	98.7	190	2	AAY42274	Chimeric protein
21	464	98.7	190	2	AAY42275	Chimeric protein
22	464	98.7	191	2	AAW42270	Human anti-androge
23	464	98.7	191	2	AAY42276	Human GH with heter
24	464	98.7	262	2	AAB71782	Human GH with heter
25	464	98.7	262	2	AAB71783	Human GH with heter
26	464	98.7	320	11	AAE42526	First G protein-li
27	464	98.7	320	11	AAE42527	First G protein-li
28	462	98.3	144	11	AAW42267	Human anti-androge
29	462	98.3	144	11	AAY42277	Human growth hormo
30	462	98.3	144	11	AAY42278	Human growth hormo
31	462	98.3	144	11	AAY42279	Human growth hormo
32	462	98.3	144	11	AAY42280	Human growth hormo
33	462	98.3	144	11	AAY42281	Human growth hormo
34	462	98.3	144	11	AAY42282	Human growth hormo
35	462	98.3	144	11	AAY42283	Human growth hormo
36	462	98.3	144	11	AAY42284	Human growth hormo
37	462	98.3	144	11	AAY42285	Human growth hormo
38	462	98.3	144	11	AAY42286	Human growth hormo
39	462	98.3	144	11	AAY42287	Human growth hormo
40	462	98.3	144	11	AAY42288	Human growth hormo
41	462	98.3	144	11	AAY42289	Human growth hormo
42	462	98.3	144	11	AAY42290	Human growth hormo
43	462	98.3	144	11	AAY42291	Human growth hormo
44	462	98.3	144	11	AAY42292	Human growth hormo
45	462	98.3	144	11	AAY42293	Human growth hormo

ALIGNMENTS

RESULT 1
AAY42267
ID AAY42267 standard; protein; 92 AA.
XX
AC AAY42267;
XX
DT 19-JAN-2000 (first entry)
XX
DE Human growth hormone (hGH) N-terminal truncated form.
XX
KW Growth hormone; hGH; pituitary hormone; acromegaly; gigantism;
KW pituitary dwarfism.
XX
OS Homo sapiens.
XX
PN WO9958562-A1.
XX
PD 18-NOV-1999;
XX
PF 12-MAY-1999; 99WO-US010468.
XX
PR 12-MAY-1998; 98US-0085104.
XX
PA (GENE) GENENTECH INC.
XX
PI Lai J;
XX
DR WPI; 2000-044222/03.
XX
PT Novel human growth hormone variant, useful for treating growth
PT disorders, comprises 1-92 amino acids of wild-type hGH.
XX

```
PS   Claim 5; Page 28; 46pp; English.
XX
CC   The sequence represents an N-terminal fragment of human growth hormone
CC   (hGH) which is a component of a chimeric protein (AAY12471) which also
CC   contains a human insulin precursor (AAY12670). The hGH portion of the
CC   chimeric protein acts as an intramolecular chaperone (IMC) for the
CC   insulin precursor, enabling it to fold correctly. A cleavable peptide
CC   linker with a C-terminal Arg residue (AAY12467) enables the hGH portion
CC   of the chimeric protein to be removed after folding has taken place.
CC   Production of recombinant human insulin via an hGH-proinsulin chimeric
CC   protein can provide human insulin with correctly linked cysteine bridges
CC   with fewer necessary procedural steps, and hence resulting in a higher
CC   yield of human insulin. The IMC sequences of the fusion proteins
CC   sequences from intracellular degradation by a microorganism host, but
CC   also promote the folding of the fused insulin precursor, facilitate the
CC   solubility of the fusion protein and decrease the intermolecular
CC   interactions among the fusion proteins. This allowing folding of the
CC   fused insulin precursor at commercially useful high concentrations. The
CC   procedural steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC   and related purification steps can thus be eliminated, along with the
CC   use of high concentrations of mercaptan or the use of hydrophobic
CC   affinity resins.
XX
SQ   Sequence  92 AA;

  Query Match             100.0%;  Score 470;  DB 20;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-54;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MPTIFLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPTIFLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 60

Qy      61 TGSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |||||||||||||||||||||||||||||||
Db      61 TGSNREETQQKSNLELLRISLLLIQSWLEPVQ 92


RESULT 2
AAW62266
ID   AAW62266 standard; protein; 154 AA.
XX
AC   AAW62266;
XX
DT   29-JUN-2000  (first entry)
XX
DE   Human anti-angiogenic peptide 2/K bla Met fragment.
XX
KW   Human; anti-angiogenic; prolactin; placental lactogen; hPL; angiogenesis;
KW   growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;
KW   placental vascularisation; pregnancy; treatment; angiogenic disease;
KW   cancer; neoplasia; malignant; arteriovenous malformation;
KW   arthritis; atherosclerotic plaque; corneal graft neovascularisation;
KW   delayed wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW   granulation; glaucoma; ocular; uveitis; fracture; osler-Weber syndrome;
KW   psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
KW   leukaemia; solid tumour; reproductive disorder; contraceptive use;
KW   detection; pre-eclampsia; intrauterine growth retardation;
KW   placental dysfunction.
XX
OS   Homo sapiens.
XX
PN   WO9950302-A1.
XX
PD   07-OCT-1999.
XX
PF   31-MAR-1999;  99WO-US006961.
XX
PR   31-MAR-1998;  98US-0080601.
XX
PA   (GELT ) GENTECH INC.
XX
PI   Martin JA, Stroman L, Taylor R, Wellner KI;

XX
DR   WPI; 1999-601762/51.
DR   N-PSDB; AAX17541.
XX
PT   New anti-angiogenic peptides - comprise N-terminal fragment of
PT   human placental lactogen, human growth hormone, growth hormone
PT   variant or human prolactin.
XX
PS   Claim 1; Page 71; 97pp; English.
XX
CC   This invention describes novel human anti-angiogenic peptides derived
CC   from 10 to 150 consecutive amino acids selected from the N-terminal end
CC   of human placental lactogen (hPL), human growth hormone (hGH), its 20kDa
CC   hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
CC   capillary endothelial cell proliferation and organisation, (ii) inhibit
CC   angiogenesis in chick chorioallantoic membrane and (iii) bind to at
CC   least one specific receptor which does not bind an intact hPL, growth
CC   hGH, hPL, prolactin or hGH-V. The invention also describes a method for
CC   diagnosing a probable abnormality of placental vascularisation during
CC   pregnancy. The peptides can be used for treating an angiogenic disease in
CC   a subject, for inhibiting tumour formation or growth in a patient, for
CC   modulating vascularisation of a patient's placenta in pregnancy. The
CC   peptides can be used for prevention or treatment of malignant and non-
CC   angiofibroma, arteriovenous malformation, arthritis such as rheumatoid
CC   arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC   delayed wound healing, proliferative retinopathy such as diabetic
CC   retinopathy, macular degeneration, granulations such as those occurring
CC   in haemarthic joints, inappropriate vascularisation of wound healing
CC   such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular
CC   tumour, uveitis, granulation fractures, osler-Weber syndrome, psoriasis,
CC   pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,
CC   Kaposi's sarcoma, trachoma, vascular adhesions, chronic inflammatory
CC   leukaemia, and reproductive disorders such as follicular and ovarian
CC   and uterine tumours. They can also be used as contraceptive agents. DNA
CC   encoding the peptides can be used in gene therapy. The measurement of
CC   abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL
CC   can be used in assays for impairment of vascular development associated
CC   with pre-eclampsia, intrauterine growth retardation, and placental
CC   dysfunction.
XX
SQ   Sequence  154 AA;

  Query Match             100.0%;  Score 470;  DB 20;  Length 154;
  Best Local Similarity   100.0%;  Pred. No. 8.6e-54;
  Matches   92;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MPTIFLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MPTIFLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTPS 60

Qy      61 TGSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |||||||||||||||||||||||||||||||
Db      61 TGSNREETQQKSNLELLRISLLLIQSWLEPVQ 92


RESULT 3
AAY12467
ID   AAY12467 standard; protein; 92 AA.
XX
AC   AAY12467;
XX
DT   10-JAN-2000  (first entry)
XX
DE   Chimeric protein fusion 5.
XX
KW   Insulin; precursor; growth hormone; chaperone; intramolecular;
KW   folding; conformation; chimeric protein; cleavable; recombinant;
KW   production; yield.
XX
OS   Synthetic.
XX
PN   WO9950302-A1.
XX
```

XX
XX
PT [illegible]
XX
RT [illegible]
XX
RL [illegible]
XX
RA (1 Y.L) [illegible]
XX
RT [illegible]
XX
RF WO [illegible]
XX
PI New human growth hormone (hGH) [illegible]
PI [illegible]
PI [illegible]
XX
PS Disclosure; pp; 22 pages; English.
XX
CC Human hGH with a 13 amino acid [illegible] of the N-terminus, a 3 amino
CC protein, which contains a thrombin cleavage site, and human insulin
CC growth factor (IGF-1 N'T) at the C-terminus. Beta N'T can be used
CC diabetes patients, especially those who do not respond to
CC recombinant human insulin. The fusion protein is produced in
CC [illegible] and is secreted into the culture medium. The fusion
CC AAN96077 B, AAP00, SC, AAP00, [illegible].
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60
[illegible]
Db 1 MFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
[illegible]
Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 1
AAP00964
[illegible] AA; 1 standard protein; [illegible] AA.
XX
AC AAP00964;
XX
[illegible]
XX
DT [illegible] (first entry)
XX
DE Human growth hormone segment.
XX
KW human growth hormone; hGH; gene therapy;
KW [illegible] insulin; IGF; diabetes; [illegible]
XX
OS Homo sapiens (human).
XX
PN JP2001046081-A.
XX
PD [illegible].
XX
PF 02-DEC-1999; 99JP-0345001.
XX
PR [illegible]
XX
PA (WARP) WARUNA PRIVAT. KK.
XX
CC WO [illegible].
CC N-terminal AAN96077.
XX
FH Key Location/Qualifiers
FT misc_feature 1..[illegible]
FT /note= "[illegible]"
XX
FH [illegible]
XX
FT [illegible]
XX
PS [illegible]

XX
CC the invention consists of a vector conta. a fusion protein which is
CC t ined by fla ting, downstream of a promoter, hGH or a deriv. (prot.
CC t ined by subtstn. of Met 14 with Leu) and a foreign protein.
CC Stability of the vector in the host is greatly increased so the
CC p tein yield is higher.
XX
SQ Sequence 192 AA;

Query Match 99.0%; Score 465; DB 2; Length 192;
Best Local Similarity 98.9%; Pred. No. 3.9e-59;
Matches 91; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
||||||||||||||||||||||||||||||
Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

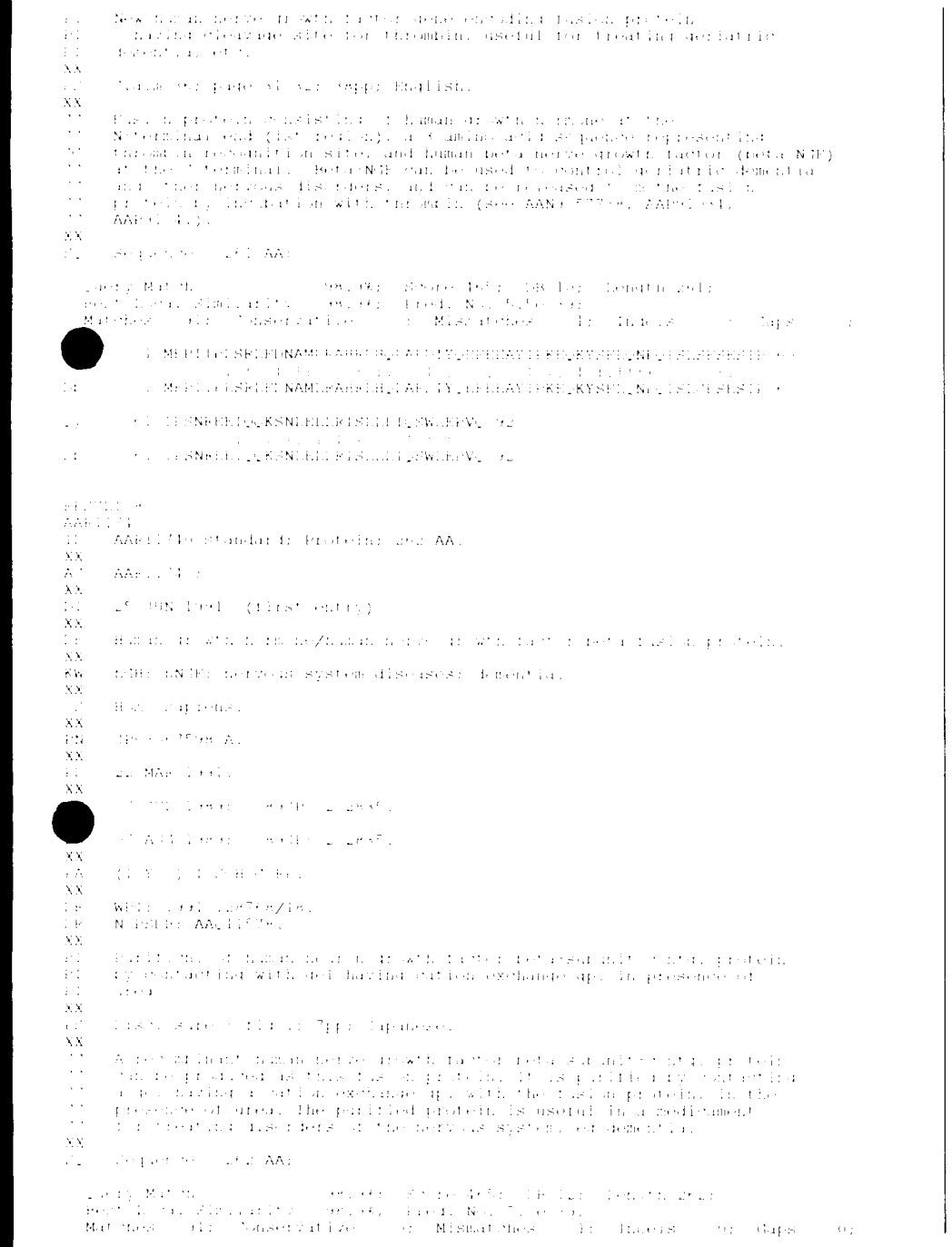
RESULT 7
AAW62294
ID AAW62294 standard; Protein; 192 AA.
XX
AC AAW62294;
XX
DT 14-NOV-1990 (first entry)
XX
DE Human anti-natriuretic peptide hGH Met-lineage.
XX
KW Human; anti-natriuretic; prolactin; placental lactogen; hGH; angiogenesis;
KW growth hormone; hGH; hGHV; capillary endothelial cell proliferation;
KW placental vascularisation; pregnancy; treatment; angiogenic disease;
KW tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;
KW arthritis; atherosclerotic plaques; corneal graft neovascularisation;
KW wound healing; proliferative retinopathy; macular degeneration; trachoma;
KW granulation; glaucoma; ocular; uveitis; fracture; osler Weber syndrome;
KW psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;
KW ovarian; leukaemia; reproductive disorder; contraception;
KW intrauterine therapy; pre-eclampsia; intrauterine growth retardation;
KW placental dysfunction.
XX
OS Homo sapiens.
XX
PN W 9005162-A1.
XX
PD 19-NOV-1990.
XX
PF 12-MAY-1990; 90WO-US02694.
XX
PR 12-MAY-1989; 89US-0350941.
XX
PA (REGE-) UNIV CALIFORNIA.
XX
PI Martial JA, Struman I, Taylor S, Weiner RI;
XX
DR WPI; 1990-345192/47.
DR N-PSDB; AAX01700.
XX
PT New anti-natriuretic peptides - comprise N-terminal fragments of
PT human placental lactogen, human growth hormone or variant or human prolactin
XX
CC Example 2; Page 69; 87pp; English.
XX
CC This invention describes a new human anti-natriuretic peptide derived
CC from 14 to 15 consecutive amino acids selected from the N-terminal end
CC of human placental lactogen (hPL), human growth hormone (hGH), growth
CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit
CC capillary endothelial cell proliferation and organisation (ii) inhibit
CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at
CC least one specific receptor which does not bind an intact hormone
CC hGH, hPL, prolactin or hGH-V. The invention also describes a method for
CC diagnosing a probable abnormality of placental vascularisation in fetal
CC pregnancy. The peptides can be used in treating an anti-social tumour in
CC a subject, for inhibiting tumour formation or growth in a patient, or for
CC modulating vascularisation of a patient's placenta, in particular. The
CC peptides can be used for preventing or treating e.g. malignant or non-
CC angiofibroma, arteriovenous malformation, arthritis such as rheumatoid
CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,
CC delayed wound healing, proliferative retinopathy such as diabetic
CC retinopathy, macular degeneration, granulations such as those occurring
CC in haemophilic joints, inappropriate vascularisation in wound healing
CC such as hypertrophic scars or Keloid scars, hemangioma, thyroid
CC tumour, uveitis, non-union fractures, Osler Weber syndrome, psoriasis,
CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,
CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic inflammation,
CC leukaemia, and reproductive disorders such as follicular and luteal cysts
CC and choriocarcinoma. They can also be used as contraceptive agents. DNA
CC encoding the peptides can be used in gene therapy. The measurement of
CC abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL
CC can be used in assays for impairment of vascular development associated
CC with pre-eclampsia, intrauterine growth retardation, and placental
CC dysfunction.
XX
SQ Sequence 192 AA;

Query Match 98.9%; Score 465; DB 2; Length 192;
Best Local Similarity 97.8%; Pred. No. 7.9e-59;
Matches 90; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MEFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MEFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP 60

Qy 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||||||||||||||||||||||||||
Db 61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 7
AAE61299
ID AAE61299 standard; protein; 201 AA.
XX
AC AAE61299;
XX
DT 14-DEC-1999 (first entry).
XX
DE Human hormone growth factor and human growth hormone fusion protein.
XX
KW Human hormone growth factor; fusion protein; thrombin;
KW psychiatric disorder; nervous disorder; human growth hormone.
XX
OS Homo sapiens (human).
XX
FH Key Location/Qualifiers
FT Active. 1..191
FT Region 141..191
FT Region 141..201
XX
PN EP-1017715-A.
XX
PD 12-JUL-2000.
XX
PF 17-FEB-1999; 99EP-0102705.
XX
PR 19-FEB-1998; 98GB-0003942.
XX
PA (FERR-) FERRING BV.
XX
PI Ohtsuka E;
XX
DR WPI; 1999-484192/41.
XX



```
XX
KW   B-cell stimulatory factor-2; interleukin-6.
XX
OS   Homo sapiens.
XX
RN   [1]
XX
RC   17-JAN-1990.
XX
RL     [JOU] 1988:    88:3:1627-1632.
XX
RL     [ATL] 1988:    88:3:1627-1632.
XX
RA   (IYJU ) IoSoH CoRP.
XX
RT   Wri; ,,, ,,,,2/c1.
RN   N IStDb; AAQ02028.
XX
PT   Expln. of human B-cell differentiation factor - from specified DNA
PT   sequence segment, by recombinant DNA technique, given protein of
PT   specified amino acid sequence.
XX
PS   Disclosure; Page 7c; 17pp; Japanese.
XX
CC   The sequence encoding this protein can be used with DNA encoding B-cell
CC   differentiation factor (IL-6) and ligated into an expression vector for
CC   prod. of a fusion protein.
CC   See also; AAB74111.
XX
SQ   Sequence   184 AA;

  Query Match             98.3%;  Score 462;  DB 11;  Length 144;
  Best Local Similarity   97.8%;  Pred. No. 7.2e-49;
  Matches  180;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MFPTIPLSELFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP  60
           || || | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFPTIPLSELFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP  60

Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ  92
           |||||||||||||||||||||||||||||||
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ  92


RESULT 12
AAP81226
ID   AAP81226 standard; protein; 262 AA.
XX
AC   AAP81226;
XX
DT   25-NOV-2002  (first entry)
XX
DE   Sequence of protein with post-translational modification site.
XX
KW   Growth hormone.
XX
OS   Synthetic.
XX
PN   JP2002272486-A.
XX
PD   13-SEP-2002.
XX
PF   20-MAR-2001;  2001JP-0079173.
XX
PR   20-MAR-2000;  2000JP-0078504.
XX
PA   (IYJU ) IoYo SoDA MFG KK.
XX
DR   WPI; 2002-722515/77.
DR   N-PSDB; AAN81605.
XX
PT   Polypeptide with a post-translational site,
PT   by culturing bacterium transformed by plasmid vector, gene
PT   segment with specified DNA sequence.
XX
PS   Claim 2(1); Page 5pp; 9pp; Japanese.
XX
CC   The polypeptide (AAP81227) with a post-translational site, and its cDNA
CC   (AAN81606) encoding it are claimed. A Met residue may be added to
CC   the N-terminal. The polypeptide acts on the bone structure, i.e. it is
CC   including humans, to promote bone growth with the polypeptide has high
CC   production rate and is easily extracted from bacterial culture medium
CC   and refined for use as a bone growth accelerator.
XX
SQ   Sequence   188 AA;

  Query Match             97.8%;  Score 460;  DB 11;  Length 188;
  Best Local Similarity   96.6%;  Pred. No. 8.6e-49;
  Matches  180;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 MFPTIPLSELFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP  60
           || || | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MFPTIPLSELFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIP  60

Qy      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ  92
           |||||||||||||||||||||||||||||||
Db      61 TPSNREETQQKSNLELLRISLLLIQSWLEPVQ  92
```

PA (SUMI) SUMITOMO SEIYAKU KK.
XX
PR Wrl: 1999-283567/24.
XX
PT A human growth hormone mutant - with equivalent activity to natural
PT human growth hormone
XX
PS Example 1; Page 5 6; 10pp; Japanese.
XX
CC The present invention describes a human growth hormone mutant in which
CC the 134th Arg and the 135th Thr are replaced respectively by Asp and Pro
CC in the 1st to the 191st amino acid sequence of natural type human 22 kDa
CC growth hormone (hGH) and which has a resistance against decomposition by
CC thrombin. The present sequence represents the natural hGH. Also
CC described are: (1) a hGH mutant in which the 134th Arg, the 135th Thr
CC and the 140th Lys are replaced respectively by Asp, Pro and Ala in the
CC amino acid sequence of natural type hGH and which has a resistance
CC against decomposition by thrombin and plasmin; and (2) a drug
CC preparation containing the above hGH mutant as the active component.
CC The mutant hGH shows an activity approximately equivalent to that of
CC natural type hGH and shows a high stability in blood and body fluid.
XX
SQ Sequence 191 AA;

Query Match 97.9%; Score 460; DB 20; Length 191;
Best Local Similarity 98.9%; Pred. No. 1.2e-38;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60
|:|||||||||:|||||||||||||||||||||||||||||||||||| ||||| ||||||
Db 1 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 60

Qy 61 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
||:|||||||||||||||||||||||:|||||||
Db 61 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 91

Search completed: February 5, 2003, 12:44:26
Job time : 68.5914 secs

Protein - protein search using sw model

Run on: [illegible] Scoring table: [illegible] Gapext [illegible]
(x10 GCUPS)
[illegible] Million cell updates/sec

Title: US-09-423-100-2
Sequence: 1 MEIIPLSKLFLNAMLRAHK.....NLLLELSLLYSWIFPV 92

Scoring Table: BLOSUM62
Gapop [illegible], Gapext [illegible]

Searching: [illegible] sequences; [illegible] residues

[illegible] matched; [illegible] scored

Minimum DB seq length:
Maximum DB seq length: [illegible]

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

The N... is the number of results predicted by chance to have a score greater than or equal to the score listed. The higher the score, the smaller the N..., and the more significant the result. These N... values are calculated using the distribution of scores for the query sequence (see Altschul, S. F. & Gish, W., Meth. in Enzymology, in press).

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4[illegible]	[illegible]	217	1	S1B2	somatotropin precursor
2	4[illegible]	[illegible]	217	2	I4754	somatotropin - rhesus
3	4[illegible]	[illegible]	217	1	S1BRV	somatotropin precursor
4	4[illegible]	[illegible]	26[illegible]	1	S1B0V2	somatotropin precursor
5	4[illegible]	[illegible]	217	2	I4741	somatotropin - rhesus
6	4[illegible]	[illegible]	217	2	I67409	growth hormone - bonnet
7	4[illegible]	[illegible]	26[illegible]	2	I67406	chorionic somatomammotropin
8	4[illegible]	[illegible]	217	2	I5[illegible]7	chorionic somatomammotropin
9	3[illegible]	[illegible]	217	1	C7BD	chorionammotropin
10	3[illegible]	[illegible]	217	2	B26445	chorionammotropin
11	3[illegible]	[illegible]	215	2	A26449	chorionammotropin
12	3[illegible]	[illegible]	216	2	B49157	somatotropin - orc
13	3[illegible]	[illegible]	1[illegible]	2	BN[illegible]4	somatotropin - sei
14	3[illegible]	[illegible]	22[illegible]	1	S1MS	somatotropin precu
15	3[illegible]	[illegible]	1[illegible]	1	S15	somatotropin - bat
16	3[illegible]	[illegible]	21[illegible]	1	S1e[illegible]	somatotropin precu
17	3[illegible]	[illegible]	21[illegible]	2	S4[illegible]e	somatotropin precu
18	3[illegible]	[illegible]	1[illegible]	2	I8[illegible]9	somatotropin - Afr
19	3[illegible]	[illegible]	21[illegible]	1	S16[illegible]	somatotropin precu
20	3[illegible]	[illegible]	21[illegible]	2	I4613[illegible]	somatotropin precu
21	3[illegible]	[illegible]	21[illegible]	2	B7[illegible]2	somatotropin precu
22	3[illegible]	[illegible]	26[illegible]	2	A47[illegible]2	somatotropin precu
23	3[illegible]	[illegible]	1[illegible]	1	A0[illegible]4	somatotropin - grp
24	3[illegible]	[illegible]	1[illegible]	2	JS042[illegible]	somatotropin - Afr
25	3[illegible]	[illegible]	21[illegible]	1	S1B[illegible]	somatotropin precu
26	3[illegible]	[illegible]	21[illegible]	1	S1SH	somatotropin precu
27	3[illegible]	[illegible]	21[illegible]	1	S1GI	somatotropin precu
28	3[illegible]	[illegible]	21[illegible]	2	S42[illegible]2	somatotropin - dog
29	3[illegible]	[illegible]	26[illegible]	2	I[illegible]4	somatotropin precu
30	3[illegible]	[illegible]	21[illegible]	2	I4[illegible]2	somatotropin precu
31	3[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
32	3[illegible]	[illegible]	21[illegible]	2	I[illegible]2	somatotropin precu
33	3[illegible]	[illegible]	21[illegible]	2	C[illegible]2	somatotropin precu
34	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
35	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
36	2[illegible]	[illegible]	21[illegible]	2	I[illegible]2	somatotropin precu
37	2[illegible]	[illegible]	21[illegible]	2	C[illegible]2	somatotropin precu
38	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
39	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
40	2[illegible]	[illegible]	21[illegible]	2	I[illegible]2	somatotropin precu
41	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
42	2[illegible]	[illegible]	21[illegible]	2	I[illegible]2	somatotropin precu
43	2[illegible]	[illegible]	21[illegible]	2	C[illegible]2	somatotropin precu
44	2[illegible]	[illegible]	21[illegible]	2	A[illegible]2	somatotropin precu
45	2[illegible]	[illegible]	21[illegible]	2	I[illegible]2	somatotropin precu

ALIGNMENTS

RESULT 1
S1B2
somatotropin precursor - rhesus
N:Alternate names: growth hormone; growth hormone precursor
N:Contains: growth hormone 5K peptide; somatotropin 22K peptide
C:Species: Homo sapiens (man)
C:Date: 24-Apr-1984 #sequence_revision 1-Jul-1987 #text_change 16-Jul-
C:Accession: A01761; A24565; A26401; A26411; A26412; A26413; A26414; A
K:Goodman, H.M.; Moore, D.D.; Seeburg, P.H.
Nucleic Acids Res. 9, 4251-4258, 1981.
A:Title: Human growth hormone cDNA sequence and mRNA structure: possible
A:Reference number: A24565; MUID:82017911; PMID:6270677
A:Accession: A24565
A:Molecule type: mRNA
A:Residues: 1-217 <GOO>
A:Cross-references: GB:V00519
A:Note: the 5' K shift was used to position cross-hybridizing fragments
K:Goodman, H.M.; Olson, M.V.; Hall, B.D.; Nature 287, 609-613, 1980
Genomics 4, 479-497, 1989
A:Title: The human growth hormone gene: nucleotide sequence, biology, and
A:Reference number: A24565; MUID:89170117
A:Accession: A24565
A:Molecule type: mRNA
A:Residues: 1-217 <GOO>
A:Cross-references: GB:M13438; NID:g178[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
A:Experimental source: [illegible]; clone [illegible]
Nucleic Acids Res. [illegible]
A:Title: Molecular [illegible] of the human growth hormone gene family
A:Reference number: A24565; MUID:[illegible]
A:Accession: A24565
A:Molecule type: mRNA
A:Residues: [illegible]
A:Cross-references: GB:V00519
A:Note: [illegible]
R:Martial, J.A.; Proc. Natl. Acad. Sci. U.S.A. [illegible]
Science 205, [illegible]
A:Title: Human growth hormone: complete coding sequence of cloned cDNA
A:Reference number: A24565; MUID:[illegible]
A:Accession: A24565
A:Molecule type: mRNA
A:Residues: [illegible]
A:Cross-references: GB:J00265; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
A:Experimental source: [illegible]
A:Note: [illegible]
R:Barta, A.; Proc. Natl. Acad. Sci. U.S.A. [illegible]
Arch. Biochem. Biophys. [illegible]
A:Title: Human placental growth hormone messenger RNA: its translation in
A:Reference number: A[illegible]; MUID:[illegible]
A:Contents: [illegible]
Eur. J. Biochem. [illegible]
Arch. Biochem. Biophys. [illegible]
A:Title: Human placental growth hormone: complete amino acid sequence
A:Reference number: A[illegible]; MUID:[illegible]
A:Accession: A[illegible]
A:Molecule type: [illegible]

A;Residues: 27-94;96-217 <LIC>
R;Niall, H.D.
Nature New Biol. 230, 90-91, 1971
A;Title: Revised primary structure for human growth hormone.
A;Reference number: A92797; MUID:71179255; PMID:5278648
A;Accession: A92797
A;Molecule type: protein
A;Residues: 27-51 <NIA>
R;Niall, H.D.; Hogan, M.L.; Sauer, R.; Rosenblum, I.Y.; Greenwood, F.C.
Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971
A;Title: Sequences of pituitary and placental human growth hormones: evolution
A;Reference number: A92779; MUID:71123968; PMID:5279528
A;Accession: A92779
A;Molecule type: protein
A;Residues: 119-120;157-159 <NI2>
R;Niall, H.D.
in Prolactin and Carcinogenesis, Proc. Fourth Tenovus Workshop Prolactin, Griffiths, K.,
A;Title: The chemistry of the human lactogenic hormones.
A;Reference number: A94427
A;Contents: annotation; somatotropin revi...
R;Rowley, L.A.; Dixon, J.S.; Li, C.H.
Int. J. Pept. Protein Res. 4, 281-287, 1972
A;Title: Sequence comparison of human pituitary growth hormone, human chorionic somatom
A;Reference number: A91761; MUID:73042028; PMID:4679454
A;Accession: A91764
A;Molecule type: protein
A;Residues: 27-217 <BEW>
R;...; Bewley, T.A.; Li, C.H.
Biochem. Biophys. Res. Commun. 92, 511-516, 1980
A;Title: The 23,000-dalton variant of human growth hormone: location of the amino acid d
A;Reference number: A90217; MUID:80149196; PMID:7356479
A;Contents: somatotropin, 20K short variant
A;Accession: A90217
A;Molecule type: protein
A;Residues: 48-57;73-80 <LEW>
R;Chapman, G.E.; Rogers, K.M.; Brittain, P.; Bradshaw, P.A.; Pate...; Turner, J.; Si
J. Biol. Chem. 256, 2395-2401, 1981
A;Title: The 20,000 molecular weight variant of human growth hormone: Preparation and s
A;Reference number: A92211; MUID:81117461; PMID:7462247
A;Contents: somatotropin, 20K short variant
A;Accession: A92211
A;Molecule type: protein
A;Residues: 27-57;74-79 <CHA>
R;Slash, R.N.P.; Seavey, B.K.; Leslie, L.J.; Labella, F.J.
J. Protein Chem. 2, 425-436, 1983
A;Title: Human growth hormone peptide 1-43: isolation from pituitary glands.
A;Reference number: A91466
A;Accession: A91466
A;Molecule type: protein
A;Residues: 27-69 <SIN>
A;Note: growth hormone 5K peptide has insulin potentiating activity; its physiological p
R;Kousso, V.M.J.; Rao, I.D.; NG, F.
Biol. Chem. Hoppe-Seyler 371, 423-431, 1990
A;Title: Identification of the aspartimide structure in a previously reported peptide.
A;Reference number: S09665; MUID:90134745; PMID:2376679
A;Accession: S09665
A;Molecule type: protein
A;Residues: 27-44,'L',46-47 <KOB>
R;de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.
Science 255, 306-312, 1992
A;Title: Human growth hormone and extracellular domain of its receptor: crystal structur
A;Reference number: A41728; MUID:92196577; PMID:1549776
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
A;Note: the structure of the complex with growth hormone receptor is described
R;Gray, G.L.; Baldridge, J.S.; McKeown, K.S.; Heyneker, H.L.; Chang, C.N.
Gene 39, 247-254, 1985
A;Title: Periplasmic production of correctly processed human growth hormone in Escherich
A;Reference number: 141126; MUID:86137393; PMID:3912261
A;Accession: 184549
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-26 <RES>
A;Cross-references: GB:M11732; NID:g178178; PID:AAA52654.1; PID:g178179

C;Comment: The gene for this hormone is transcribed only in somatotrophic cells of th
C;Comment: About 90% of somatotropin is the 22K long form.
C;Genetics:
A;Gene: GDB:GH1
A;Cross-reference...; GDB:119682; OMIM:139250
A;Map position: 17q23.1-17q24.3
A;Introns: 4/...; 57/...; 97/...; 155/...
C;Superfamily: prolactin
C;Keywords: alternative splicing; hormone; pituitary
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-217/Product: somatotropin 1, 2,3; 5 <...> #status experimental <EXP>
F;27-66/Product: growth hormone 5K peptide #status experimental <5KP>
F;27-57,74-217/Product: somatotropin 1, short form #status experimental <SHO>
F;79-191,208-217/Disulfide bonds: #status experimental

  Query Match 97.9%; Score 460; DB 1; Length 217;
  Best Local Similarity 98.9%; Pred. No. 4.4e-42;
  Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||.|
Db 27 FPTIPLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAP 86

Qy 62 PGNKEEAQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||||||||||||||||| ||||||||
Db 87 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 117

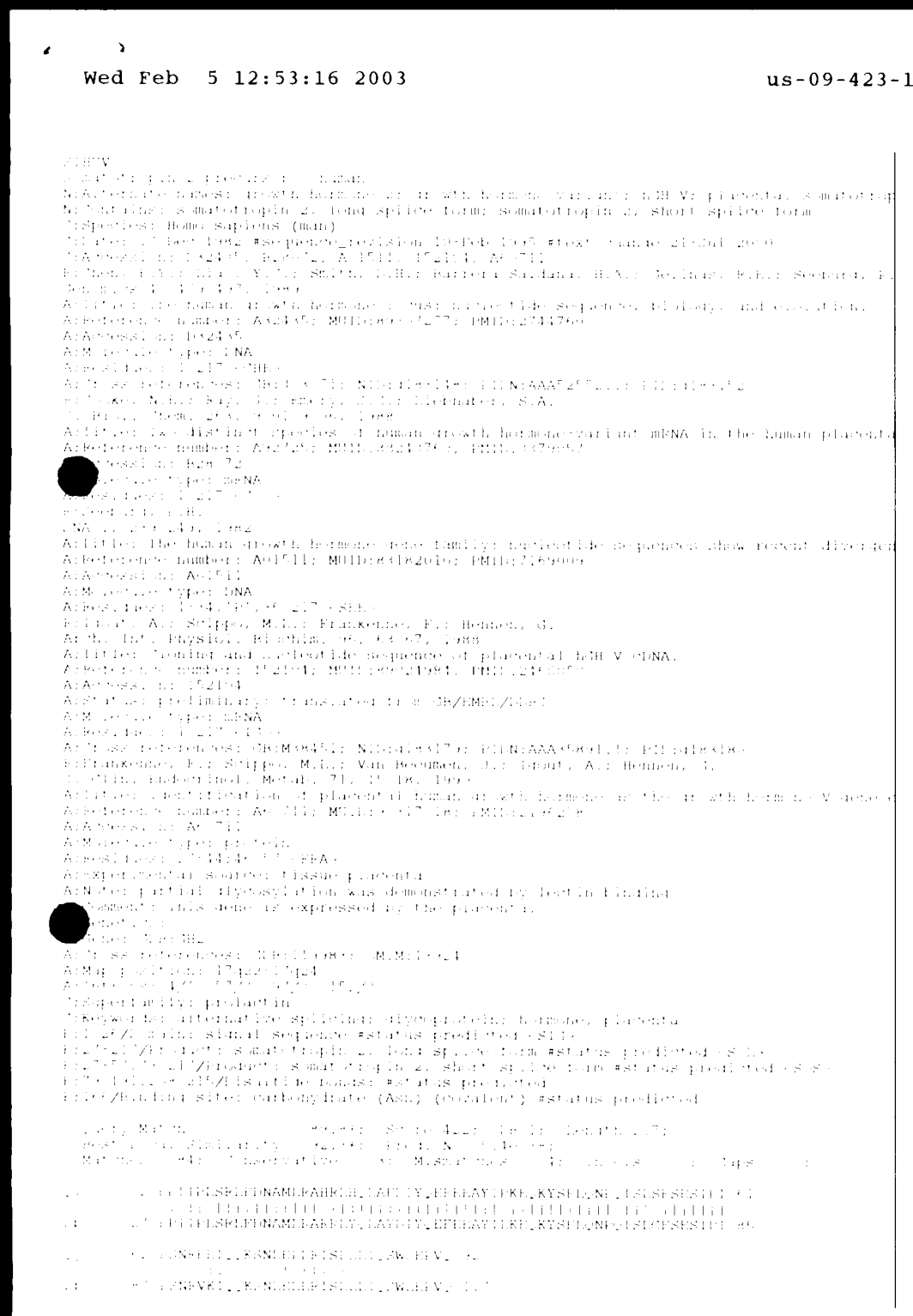
RESULT 2
167410
somatotropin - rhesus macaque
N;Alternate names: growth hormone
C;Species: Macaca mulatta (rhesus macaque)
C;Date: 31-May-1996 #sequence revision 31-May-1996 #text change 16-Jul-1999
C;Accession: 167410; A05094
R;Cubas, P.A.; Dorairaj, M.; Fiedler, J.M.; Fowler, P.B.
Endocrinology 133, 1744-1752, 1993
A;Title: Cloning of two species of chorionic somatomammotropin related complemen
A;Reference number: 153262; MUID:94009224; PMID:8103612
A;Accession: 167410
A;Status: translated from GenBank/EMBL
A;Molecule type: mRNA
A;Residues: 1-217 <RES>
A;Cross-references: GB:L16556; NID:g425115; PID:AAA36642.1; PID:g425116
R;Li, C.H.; Chung, D.; Kahn, H.W.; Stein, S.
Arch. Biochem. Biophys. 245, 287-291, 1986
A;Title: The primary structure of monkey pituitary growth hormone.
A;Reference number: A05094; MUID:86129460; PMID:3080969
A;Accession: A05094
A;Molecule type: protein
A;Residues: 27-99,'Q',101-178,'D',180-217 <LIC>
A;Note: the monkey species is not identified in the reference
R;Raben, M.S.
Science 125, 883-884, 1957
A;Title: Preparation of growth hormone from pituitaries of man and monkey.
A;Reference number: A14774
A;Contents: annotation; identification of source organism
C;Superfamily: prolactin

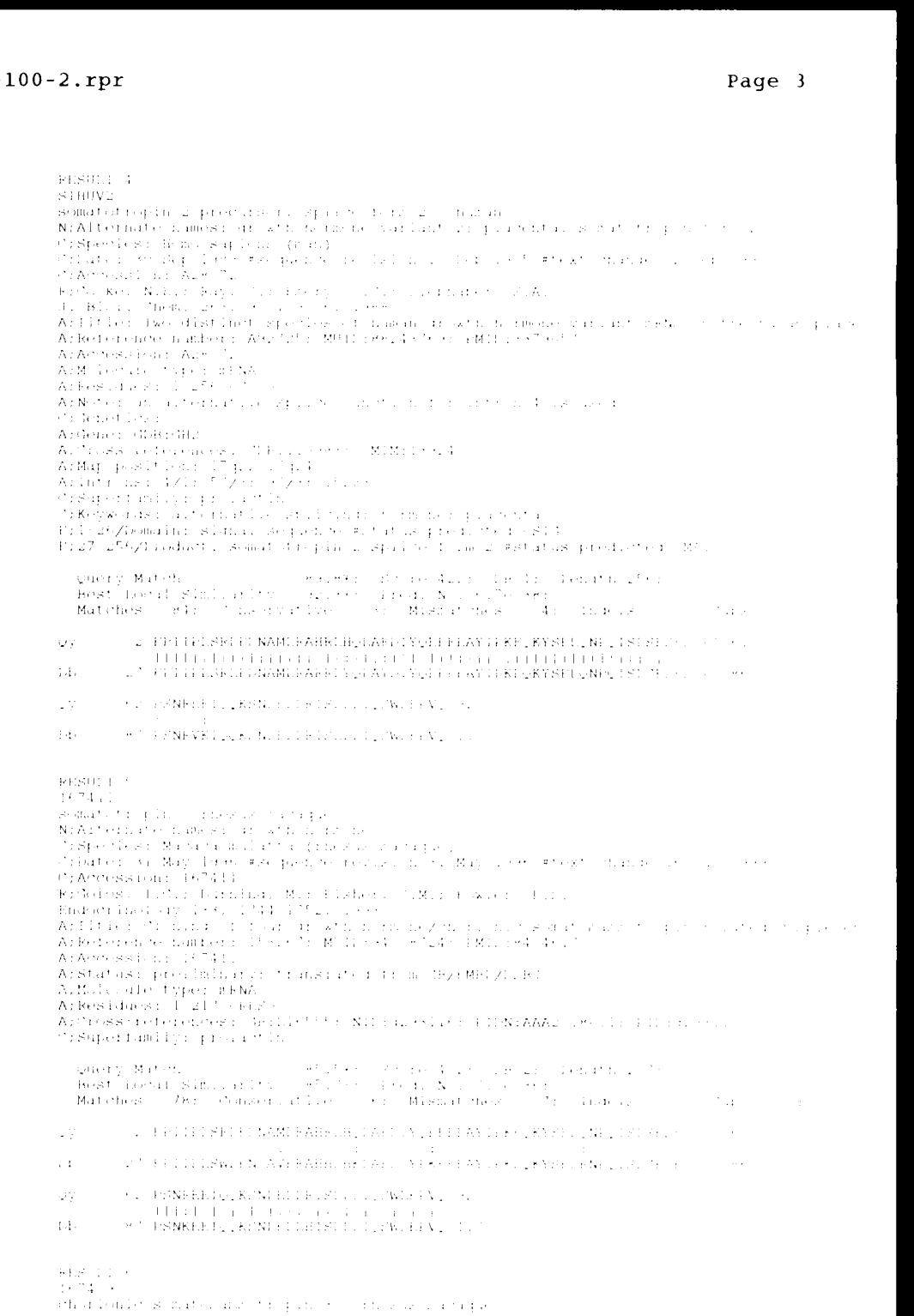
  Query Match 97.9%; Score 460; DB 2; Length 217;
  Best Local Similarity 98.9%; Pred. No. 4.4e-42;
  Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
||||||||||||||||||||| |||||||||||||||||||||||||||||| ||||.|
Db 27 FPTIPLSGLFANAVLRAQHLHQLAADTFKEFERTYIPEGQRYSIQNTQVAFCFSETIPAP 86

Qy 62 PGNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
|||||||||||||||||||||| ||||||||
Db 87 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 117

NucBase -





A;Accession: B49156
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-216 <SOU>
A;Cross-references: GB:M62292, NID:g179155, PIDN:AAB29360.1, PID:g179156
A;Note: sequence extracted from NCBI backbone (NCBIN:95219, NCBIP:95220)
C;Superfamily: prolactin

Query Match 66.1%; Score 310.5; DB 2; Length 216;
Best Local Similarity 67.0%; Pred. No. 5.4e-26;
Matches 61; Conservative 12; Mismatches 14; Indels 1; Gaps 1;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
Db 27 FPAMPLSNLFANAVLRAQHLAQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 86

Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 87 PTGKEEAQQRTDMELLRFSLLLIQSWLGPVQ 116

RESULT 13
PN0140
somatotropin - sei whale
N;Alternate names: growth hormone
C;Species: Balaenoptera borealis (sei whale)
C;Date: 07-May-1991 #sequence_revision 07-May-1991 #text_change 07-May-1999
C;Accession: PN0140
R;Yudaev, N.A.; Ilyin, V.S.; Bolt, A.A.; Filippova, L.A.
Biokhimiia 47, 1059-1069, 1982
A;Title: Amino acid sequence of sei whale somatotropin.
A;Reference number: PN0140; MUID:83030569; PMID:7115011
A;Accession: PN0140
A;Molecule type: protein
A;Residues: 1-190 <YUD>
A;Note: article in Russian with English abstract
C;Superfamily: prolactin
C;Keywords: growth factor; hormone
F;52-164,188 158/Disulfide bonds: #status predicted

Query Match 65.4%; Score 307.5; DB 2; Length 190;
Best Local Similarity 67.0%; Pred. No. 9.8e-26;
Matches 61; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Qy 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
Db 1 FPAMSLSNLFANAVLRAQHLHQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 60

Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 61 PANKDEAQQKSDVELLRFSLLLIQSWLGPVQ 90

RESULT 14
SIMS
somatotropin precursor - mouse
N;Alternate names: growth hormone
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 28-May-1999
C;Accession: B29911
R;Linzer, D.I.H.; Talamantes, F.
J. Biol. Chem. 260, 9574-9579, 1985
A;Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression
A;Reference number: A29911; MUID:85261752; PMID:2991752
A;Accession: B29911
A;Molecule type: mRNA
A;Residues: 1-216 <LIN>
A;Cross-references: GB:X02691, GB:K02232, NID:g71457, PIDN:CAA26407.1, PID:g71458
C;Superfamily: prolactin
C;Keywords: anterior pituitary; growth factor; hormone
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-216/Product: somatotropin #status predicted <SIN>
F;78 In;206 214/Disulfide bonds: #status predicted

Query Match 64.8%; Score 304.5; DB 1; Length 216;
Best Local Similarity 64.8%; Pred. No. 2.4e-25;
Matches 59; Conservative 4; Mismatches 17; Indels 1; Gaps 1;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
Db 27 FPAMPLSNLFANAVLRAQHLAQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 86

Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 87 PTGKEEAQQRTDMELLRFSLLLIQSWLGPVQ 116

RESULT 15
SIHO
somatotropin - horse
N;Alternate names: growth hormone
C;Species: Equus caballus (domestic horse)
C;Date: 15-Jul-1981 #sequence_revision 15-Jul-1981 #text_change 28-Apr-1999
C;Accession: A91772; A91795; A91383; A90240; A01514
R;Zakin, M.M.; Puskus, E.; Lanoton, A.A.; Ferrara, P.; Santome, J.A.; Bellacha, J.M.
Int. J. Pept. Protein Res. 8, 435-444, 1976
A;Title: Primary structure of equine growth hormone.
A;Reference number: A91772; MUID:77005410; PMID:965151
A;Accession: A91772
A;Molecule type: protein
A;Residues: 1-190 <ZAK>
R;Zakin, M.M.; Puskus, E.; Bellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 34, 353-355, 1973
A;Title: The amino acid sequence of equine growth hormone.
A;Reference number: A91795; MUID:73220562; PMID:4747849
A;Accession: A91795
A;Molecule type: protein
A;Residues: 1-190 <ZA2>
R;Zakin, M.M.; Puskus, E.; Bellacha, J.M.; Paladini, A.C.; Santome, J.A.
FEBS Lett. 25, 77-82, 1972
A;Title: Amino acid sequence around the cystine residues in equine growth hormone.
A;Reference number: A91383
A;Accession: A91383
A;Molecule type: protein
A;Residues: 42-69;157-190 <ZA1>
R;Oliver, L.; Hartree, A.S.
Biochem. J. 109, 19-24, 1968
A;Title: Amino acid sequence of equine growth hormone.
A;Reference number: A90240; MUID:68068050; PMID:4876100
A;Accession: A90240
A;Molecule type: protein
A;Residues: 176-190 <OLI>
C;Superfamily: prolactin
C;Keywords: hormone; pituitary
F;52-163,188 188/Disulfide bonds: #status experimental

Query Match 64.4%; Score 302.5; DB 1; Length 190;
Best Local Similarity 64.8%; Pred. No. 3.4e-25;
Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

Qy 2 FPTIPLSKLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIPT 61
Db 1 FPAMSLSNLFANAVLRAQHLHQLAADTFKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 60

Qy 62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 61 PANKDEAQQKSDVELLRFSLLLIQSWLGPVQ 90

Search completed: February 5, 2003, 12:45:50
Job time : 17.1605 secs

Title: US-09-423-100-2
Perfect score: 470

Database : SwissProt_40:*

Prob. N.G. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	94	20.0	217	1	SWMA_HUMAN	P...all homo sapien
2	46	9.8	217	1	SWMA_MACMU	P...ma macaca mula
3	44	9.4	217	1	SWMA_PANTR	P...756 pan troglod
4	44	9.4	217	1	SWMA_SAIBO	P...43 saimiri bol
5	43	9.1	217	1	SWMA_CALJA	...ontes calithrix
6	37	7.9	217	1	SWMA_PANIE	P...57 pan paniscu
7	122	26.0	217	1	SWMA_HUMAN	P...42 homo sapien
8	40	8.5	217	1	SWMA_MACMU	P...72 macaca mula
9	40	8.5	217	1	SWMA_HUMAN	P...42 homo sapien
10	39	8.3	217	1	SWMA_MESAU	P...76 mesocricet
11	39	8.3	190	1	SWMA_PAGB	P...92 balaenopter
12	38	8.1	217	1	SWMA_MUSMU	P...890 mus musculu
13	37	7.9	217	1	SWMA_RABIT	P...16 sylvilagus
14	37	7.9	217	1	SWMA_PAPIO	P...67 papio hamad
15	37	7.9	217	1	SWMA_PAN	P...214 pan troglo
16	37	7.9	217	1	SWMA_CALJA	...ontes calithrix
17	37	7.9	217	1	SWMA_MYOLU	...mmala myotis lu
18	37	7.9	217	1	SWMA_LEPEA	P...892 lobodon ca
19	37	7.9	217	1	SWMA_CANFA	P...711 canis fami
20	37	7.9	217	1	SWMA_FELCA	P...404 felis silv
21	37	7.9	217	1	SWMA_PIG	P...148 sus scrofa
22	37	7.9	217	1	SWMA_MUSVI	P...75 mustela vis
23	37	7.9	217	1	SWMA_LAMLA	P...94 lama glama
24	37	7.9	217	1	SWMA_VICVI	P...70 vicugna vic
25	37	7.9	217	1	SWMA_MESAU	P...11 mesocricet
26	37	7.9	217	1	SWMA_ERIEU	P...63 erinaceus
27	37	7.9	217	1	SWMA_BOVIN	P...16 bos taurus
28	37	7.9	217	1	SWMA_TUPBE	P...47 tupaia bel
29	37	7.9	217	1	SWMA_SHEEP	P...247 ovis aries
30	37	7.9	217	1	SWMA_RUBPI	P...95 rubeola ru
31	37	7.9	217	1	SWMA_MELCA	P...77 meleagris
32	37	7.9	217	1	SWMA_MICCA	P...49 macaca cas
33	37	7.9	217	1	SWMA_SUSSA	P...48 suncus sua
34	37	7.9	217	1	SWMA_MACFA	P...75 macaca fas
35	37	7.9	217	1	SWMA_MACMU	P...72 macaca mula
36	37	7.9	217	1	SWMA_MACNE	P...72 macaca nemes
37	37	7.9	217	1	SWMA_MACFA	P...72 macaca fas
38	37	7.9	217	1	SWMA_MACFA	P...72 macaca fas
39	37	7.9	217	1	SWMA_MACFA	P...72 macaca fas
40	37	7.9	217	1	SWMA_XENLA	P...07 xenopus lae
41	37	7.9	217	1	SWMA_PANTR	P...756 pan troglod
42	37	7.9	217	1	SWMA_BUFMA	P...43 bufo marina
43	37	7.9	217	1	SWMA_PIG	P...148 sus scrofa
44	37	7.9	217	1	SWMA_MESAU	P...76 mesocricet
45	37	7.9	217	1	SWMA_MACFA	P...72 macaca fas
46	37	7.9	217	1	SWMA_PAPAN	P...67 papio anub

ALIGNMENTS

RESULT 1
SWMA_HUMAN
ID SWMA_HUMAN STANDARD; PRT; 217 AA.
AC P...11; Q...27; Q...28; Q...29; Q...30; Q...31;
DT 01-...-1995 (Rel. 31, Created)
DT 01-MAR-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2001 (Rel. 39, Last annotation update)
DE Somatotropin precursor (Growth hormone) (GH) (GH N) (Pituitary growth
DE hormone) (Growth hormone 1).
GN GH1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; ...
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H...
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1,...
RX MEDLINE=80...; PubMed=6...;
RA Roskam W., Reddy E.;
RT "Molecular cloning and characterization of the human ...
RT Nucleic Acids Res. ...
RL Nucleic Acids Res. ...
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM ...).
RX MEDLINE=...; PubMed=...;
RA Bell G.I., Merryweather J.P., Sanchez-Pescador R., ...
RT "Sequence of a human growth hormone cDNA.";
RL Nucleic Acids Res. ...
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA Seeburg P.H.;
RT "The human growth hormone gene family: structure and ...
RT 2...
RL DNA ...
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA Goodman H.M., Bell G.I., Crea R., Hallewell R.A.;
RT "Structure of the human growth hormone gene.";
RL Nature ...
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=...; PubMed=...;
RA "Human DNA sequences from a variant growth hormone ...
RT "...";
RL ...
RN [6]
RP SEQUENCE FROM N.A. (...)
RP ...
RA ...
RT "A novel gene expressed in human pituitary and ...";
RL Submitted (...) to the EMBL/GenBank/DDBJ databases.

```
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM 4).
RC   TISSUE=Liver;
RX   MEDLINE=94345512; PubMed=7520754;
RA   Wu S.M., Bai C., Sun Y., Song H.-F., Desai Y., Jeay S., Hwang D., Bell G.I.,
RA   Leung D.W.;
RT   "Characterization of the human growth hormone-releasing factor receptor
RT   gene.";
RL   Endocrinology 135:1910-1917(1994).
RN   [8]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94012361; PubMed=7961261;
RA   Duan C., Baumbach W.R., McKeown K.S., Beckler R.L., Clark C.N.;
RT   "Periplasmic production of correctly processed human growth hormone in
RT   Escherichia coli: natural and bacterial signal sequences are
RT   interchangeable.";
RL   Gene 140:87-92(1994).
RN   [9]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93289302; PubMed=7686884;
RA   Li C.H., Dixon J.S., Liu W.K.;
RT   "Human pituitary growth hormone XIX. The primary structure of the
RT   28-kDa hGH.";
RL   Arch. Biochem. Biophys. 300:337-341(1993).
RN   [10]
RP   SEQUENCE OF 27-217, AND REVISIONS.
RX   MEDLINE=72114935; PubMed=5140027;
RA   Li C.H., Dixon J.S.;
RT   "Human pituitary growth hormone. XX. The primary structure of the
RT   hormone.";
RL   Arch. Biochem. Biophys. 151:334-339(1972).
RN   [11]
RP   REVISION.
RX   MEDLINE=75056244; PubMed=4170454;
RA   Bewley T.A., Dixon J.S., Li C.H.;
RT   "Sequence comparison of human pituitary growth hormone, human
RT   chorionic somatomammotropin, and ovine pituitary growth and
RT   lactogenic hormones.";
RL   Int. J. Pept. Protein Res. 4:281-287(1972).
RN   [12]
RP   SEQUENCE OF 1-7 AND 162-191.
RX   MEDLINE=71167065; PubMed=5279048;
RA   Niall H.D.;
RT   "Revised primary structure for human growth hormone.";
RL   Nature New Biol. 230:90-91(1971).
RN   [13]
RP   REVISIONS IN 120 AND 167-176.
RX   MEDLINE=71159059; PubMed=5279528;
RA   Niall H.D., Keutmann H.T., Sauer R., Rosenblum I.Y., Greenwood F.C.;
RT   "Sequences of pituitary and placental lactogenic and growth hormones:
RT   evolution from a primordial peptide by gene reduplication.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).
RN   [14]
RP   REVISION.
RA   Niall H.D.;
RT   "The chemistry of the human lactogenic hormones.";
RT   (In) Talmage R.V. (eds.);
RL   Proc. 1st and 1st Endocrinology Workshop Research Conf.,
RL   pp.15-27, Alpha Omega Alpha Press, Cardiff (1972).
RN   [15]
RP   SEQUENCE OF 1-7 (ISOFORM 2).
RX   MEDLINE=81115013; PubMed=7012224;
RA   Lingappa B.K., Lingappa K.M., Reddanna P., Bradshaw R.A., Bates R.W.,
RA   Lippoldt R., Cary P.D., Crane-Robinson C.;
RT   "The 20,000-molecular weight variant of human growth hormone.
RT   Preparation and some physical and chemical properties.";
RL   J. Biol. Chem. 256:1256-1264(1981).
RN   [16]
RP   SEQUENCE OF 50-96 (ISOFORM 3).
RX   MEDLINE=94015196; PubMed=7525647;
RA   Lewis U.J., Brewski L.E., Lewis L.G.;
RT   "The 20,000 dalton variant of human growth hormone: location of the
RT   amino acid deletions.";
RL   Biochem. Biophys. Res. Commun. 202:1-6(1994).
RN   [17]
RP   DEAMIDATION AT GLN-16 AND ASN-179.
RX   MEDLINE=82052967; PubMed=7023740;
RA   Geise R.J., Adair B.M., Wittliff J.L., Seavey B.K.;
RT   "Altered physicochemical properties of human growth hormone in vitro
RT   deamidation.";
RL   J. Biol. Chem. 256:13147-13154(1981).
RN   [18]
RP   REVIEW.
RX   MEDLINE=93321812; PubMed=8359451;
RA   Baumann G.;
RT   "Growth hormone heterogeneity: genes, isohormones, variants, and
RT   binding proteins.";
RL   Endocr. Rev. 14:424-442(1993).
RN   [19]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=88160755; PubMed=3131713;
RA   Cohen F.E., Kuntz I.D.;
RT   "Prediction of the three-dimensional structure of human growth
RT   hormone.";
RL   Proteins 2:162-166(1987).
RN   [20]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX   MEDLINE=94264799; PubMed=8197138;
RA   de Vos A.M., Ultsch M., Kossiakoff A.A.;
RT   "Human growth hormone and extracellular domain of its receptor:
RT   crystal structure of the complex.";
RL   Science 255:306-312(1992).
RN   [21]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=95079462; PubMed=7984111;
RA   Somers W., Ultsch M., de Vos A.M., Kossiakoff A.A.;
RT   "The X-ray structure of a growth hormone-prolactin receptor complex.";
RL   Nature 372:478-481(1994).
RN   [22]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RA   Chantalat L., Chirgadze N.Y., Jones N., Korber F., Navaza J.,
RA   Pavlovsky A.G., Wlodawer A.;
RT   "The crystal structure of wild-type growth hormone at 2.5 A
RT   resolution.";
RL   Protein Sci. 3:71-76(1994).
RN   [23]
RP   X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX   MEDLINE=97114025; PubMed=8943276;
RA   Sundstrom M., Lundqvist T., Rodin A., Edwards B.F.M., Norstedt G.;
RT   "Crystal structure of an oncogenic mutant of human growth hormone
RT   G120R, in complex with its receptor at 2.5 A resolution.";
RL   J. Biol. Chem. 271:32197-32203(1996).
CC   -!- FUNCTION: Plays an important role in growth control. It is a
CC       note in stimulating growth which is essential for the first
CC       other tissues to secrete IGF-I. It also stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBUNIT: Monomer, dimer, trimer, tetramer and pentamer. Exists in
CC       linked or non-covalent associations that are heterogeneous and
CC       heterogeneric associations that may form a complex with ALS,
CC       CHBP or with the cognate macromolecular complex.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- ALTERNATIVE PRODUCTS: At least 4 isoforms, 1/2a, 1/2b, (20 kDa form),
CC       2/3, 22 kDa variant, 4 and 5 are produced by alternative
CC       splicing.
CC   -!- DISEASE: DEFECTS IN GHR ARE A CAUSE OF PITUITARY DWARFISM , ANY
CC       IV.
CC   -!- PHARMACEUTICAL: Available under the names Nutropin or rec-GH
CC       (Genentech), Norditropin (Novo Nordisk), Saizen (Serono),
CC       Genotropin (Pharmacia & Upjohn), Humatrope (Eli Lilly) and Silvestrin (Sanofi Winthrop).
CC       Used for the treatment of growth hormone deficiency and for
CC       Turner's syndrome.
```

```
PN   1.
PI   Sequence 217 AA;
PA   Pearse A.G., Espiner E.A., Harris A.G., Butterfield B.;
RT   "Localization of the growth hormone gene in human
RT   acromegalic pituitaries.";
RL   Submitted (APR-1987) to the EMBL/GenBank/DDBJ databases.
CC   !!SEQUENCE Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   !! SEQUENCE 217 AA;  24 kD;  MW:  25924A2C8F4B1E9D CRC64;

  Query Match             91.5%;  Score 434;  DB 1;  Length 217;
  Best Local Similarity   96.4%;  Pred. No. 1.5e-88;
  Matches  185;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAY 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MATGSRTSLLLAFGLLCLPWLQEGSAFPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAY 60

Qy     61 IPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQ 117
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 IPKEQKYSFLQNPQTSLCFSESIPTPSNREETQQKSNLELLRISLLLIQSWLEPVQ 117


RESULT 2
SOMA_HUMAN
ID   SOMA_HUMAN     STANDARD;      PRT;   217 AA.
AC   P01241; P09467;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GH-V) (Placental-specific growth
DE   hormone) (Growth hormone 2).
GN   GH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=84124010; PubMed=6607132;
RA   Owerbach D.;
RT   "Structure of the human growth hormone gene family: evolution of
RT   polymorphic and pseudogene sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:1382-1385(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87124010; PubMed=3027473;
RA   Owerbach D., Lee P.D.K., Rutter W.J., Bell G.I.;
RT   "Human GH-variant gene structure, evolution, and expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:1526-1529(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88
RA   Cooke N.E., Baxter J.D.;
RT   "Structure of the human growth hormone gene cluster: relationship
RT   between normal and abnormal genes.";
RL   Submitted (AUG-1987) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88
RA   Cooke N.E., Liu E., Baxter J.D., Barrera-Saldana H.A.;
RA   Goodman H.M.;
RT   "The human growth hormone gene family: structure and evolution.";
RL   Genomics 4:17-27(1988).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Fiddes J.C.;
RT   "Sequence of cDNA clone encoding human placental growth hormone.";
RL   Submitted (JAN-1989) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   REVIEW.
RX   MEDLINE=
RA   Baumann G.;
RT   "Growth hormone heterogeneity: genes, isohormones, variants and
RT   binding proteins.";
RL   Endocr. Rev. 12:424-449(1991).
CC   !! SUBUNIT: Binds as an antiparallel dimer to two molecules of the
CC       hormone binding prolactin receptor. The first binding site on
CC       the hormone is of high affinity and determines species specificity
CC       (primate GH binds primate but not rodent prolactin receptor). The
CC       second binding site is of low affinity.
CC   !! SUBCELLULAR LOCATION: Secreted.
CC   !! ALTERNATIVE PRODUCTS: Event=Alternative splicing; Named isoforms=2;
CC       Variety 1: GH-N; Described above;
CC   !! TISSUE SPECIFICITY: Expressed in the placenta.
CC   !! SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; K00470; AAA36019.1; -.
DR   EMBL; J03756; AAB25442.1; -.
DR   EMBL; J03756; AAB25443.1; -.
DR   EMBL; M36547; AAA36019.1; JOINED.
DR   EMBL; M36547; AAA36020.1; JOINED.
DR   EMBL; M
DR   EMBL; M
DR   PIR; A
DR   PIR; S
DR   PIR; I
DR   HSSP; P01241; 1AZU.
DR   Genew; HGNC:4262; GH2.
DR   MIM; 150
DR   InterPro; IPR
DR   Pfam; PF
DR   PRINTS; PR
DR   PROSITE; PS
DR   PROSITE; PS
KW   Hormone; Secreted; Signal; Alternative splicing;
FT   SIGNAL       1     26       BY SIMILARITY.
```

```
FT   CHAIN        27    217       GROWTH HORMONE VARIANT.
FT   DISULFID     53    165       BY SIMILARITY.
FT   DISULFID     182   189       BY SIMILARITY.
FT   VARIANT      69    69        N->K,NKEE (IN 'NA'. . .) (1 ENTRIES).
FT   VARSPLIC     153   217       MISSING (IN GROWTH HORMONE VARIANT).
FT                                KMERKVEIFLEIVQPSTVFSQVSNDLVLGHSLGLSF
FT                                LASRNAREKHPDVYLSSYVALKGNLRPVSRTLYNPEN
FT                                PQTEDQKKMNQLEDFLKLVVMSIRSMSKMSPLGLPE
FT                                QGLKSVLDEEV (IN REF. 2).
FT   CONFLICT     19    19        T -> I (IN REF. 2).
SQ   SEQUENCE   217 AA;  24221 MW;  7B9247D6F8A2EF0C CRC64;

  Query Match             89.8%;  Score 422;  DB 1;  Length 217;
  Best Local Similarity   92.6%;  Pred. No. 6e-56;
  Matches  201;  Conservative    5;  Mismatches    4;  Indels    5;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           |||||||||||||||||| |||||:|||||||| ||||||||||||||| |||||||
Db      27 FPTIPLSKLFDNAMLRAKRLYQLAYDTYQEFEEAYIPKDQKYSFLQNPQTSLCFSESIPT 86

Qy      62 FSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |:|:|:|:|:|||||| |||||||||||||
Db      87 FSNKVKTQQKSNLELLHISLLLIQSWLEPVQ 117


RESULT 2
S_ML_MAMM
ID   S_ML_MAMM      STANDARD;      PRT:   217 AA.
AC   P_____; Q_____;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Growth hormone variant precursor (GHV) (Placental specific 1 with
DE   insulin) (Growth hormone 2).
GN   GHL.
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Golos T.G.;
RT   Submitted (JAN 1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=95008724; PubMed=8016117;
RA   Golos T.G., Durning M., Fisher J.M., Fowler P.B.;
RT   "Placental and brain growth hormone/chorionic somatomammotropin variant
RT   expression in the rhesus monkey placenta.";
RL   Biol. Reprod. 51:1244-1252(1994).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       milk, fatty acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Expressed in the placenta.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ---------- ---- ------- ------- ------- ------- ------- ------- --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---- ------- ------- ------- ------- ------- ------- --
DR   EMBL; M_____; AAA_____.1; -.
DR   EMBL; L16555; AAA20190.1; -.
DR   BSSP; P01241; IHGH.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; Somatotropin; 1.
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN; 1.
DR   PROTEIN; PROTOMAP; SOMATOTROPIN; 1.
KW   Hormone; Placenta; Signal; Glycoprotein.
FT   SIGNAL       1     26        BY SIMILARITY.
FT   CHAIN        27    217       GROWTH HORMONE VARIANT.
FT   DISULFID     53    165       BY SIMILARITY.
FT   DISULFID     182   189       BY SIMILARITY.
FT   CONFLICT     57    57        L -> F (IN REF. 2).
FT   CONFLICT     152   152       E -> G (IN REF. 2).
SQ   SEQUENCE   217 AA;  24221 MW;  7B9247D6F8A2EF0C CRC64;

  Query Match             89.8%;  Score 422;  DB 1;  Length 217;
  Best Local Similarity   92.6%;  Pred. No. 6e-56;
  Matches  201;  Conservative    5;  Mismatches    4;  Indels    5;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           |||||||||||||||||| |||||:|||||||| ||||||||||||||| |||||||
Db      27 FPTIPLSKLFDNAMLRAKRLYQLAYDTYQEFEEAYIPKDQKYSFLQNPQTSLCFSESIPT 86

Qy      62 FSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |:|:|:|:|:|||||| |||||||||||||
Db      87 FSNKVKTQQKSNLELLHISLLLIQSWLEPVQ 117


RESULT 3
PEL_HUMAN
ID   PEL_HUMAN      STANDARD;      PRT:   217 AA.
AC   P01241;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-APR-1990 (Rel. 07, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Lactogen precursor (Chorionmammotropin) (Chorionic somatomammotropin).
GN   CSH1 AND CSH2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (GENE CSH1).
RX   MEDLINE=85xxxx26; PubMed=3xxx702;
RA   Selby M.J., Barta A., Boime I., Roe B., Sussman H., Eberhardt N.;
RT   "Analysis of a full human gene encoding placental lactogen.";
RL   Int Symp Endocrinol Prod Ovine Bovine Porcine :1-1(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (GENE CSH1).
RX   MEDLINE=87xxxx25; PubMed=3xxx5xx;
RA   Barta A., Richards R.I., Baxter J.D., Shine J., Shine J.;
RT   "Evolution of genes encoding human growth hormone, chorionic somatomammotropin
RT   and placental lactogen.";
RL   DNA 6:539-546(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=xxxxxxx5; PubMed=xxxxxxx;
RA   Barrera-Saldana H.A., Seeburg P.H., Rinderknecht E.;
RT   "Two structurally different genes produce the same secreted human
RT   placental lactogen hormone.";
RL   J. Biol. Chem. 260:8307-8313(1985).
RN   [4]
RP   SEQUENCE FROM N.A. (GENES CSH1 AND CSH2).
RX   MEDLINE=xxxxxxx7; PubMed=2xxxxx6;
RA   Chen E.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gelinas R.E.,
RA   Seeburg P.H.;
RT   "Human growth hormone and placental lactogen: protein sequence, nucleic
RT   acid structure and gene organization.";
RL   Genomics 4:479-497(1989).
RN   [5]
RP   SEQUENCE.
```

```
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00466; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    216       SOMATOTROPIN.
FT   DISULFID     76    164       BY SIMILARITY.
FT   DISULFID    206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24696 MW;  3B6C9D12AB6F1166 CRC64;

  Query Match             66.1%;  Score 410.5;  DB 1;  Length 216;
  Best Local Similarity   67.0%;  Pred. No. 6e-26;
  Matches   61;  Conservative   14;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
           ||  ||| ||  |||||  ||||| |||:|||  |||:  |:|  |||  ||  ||:  |||:||
Db      27 FPAMPLSNLFANAVLRAQHLAQLAADTFKEFERTYIPEGQRYSIQNAQAAFCFSETIPAP 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |:|:|| |||:|||| |||||||||| |||
Db      87 PGKEEAQQRSDMELLRFSLLLIQSWLGPVQ 116


RESULT 11
SOMA_BALAC
ID   SOMA_BALAC     STANDARD;      PRT;   216 AA.
AC   P38992;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin (Growth hormone).
GN   GH1.
OS   Balaenoptera borealis (Sei whale).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC   Balaenopteridae; Balaenoptera.
OX   NCBI_TaxID=9773;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=83022566; PubMed=7115013;
RA   Yudaev N.A., Fankov Y.A., Evlatov A.A., Ostapova I.A.;
RT   "Amino acid sequence of soluble somatotropin.";
RL   Biokhimiia 47:1059-1069(1982).
RN   [2]
RP   PRELIMINARY PARTIAL SEQUENCE.
RA   Ostapova I.A., Soloiov A.A., Lazk  Y.S.;
RT   "Structural studies of tryptic peptides from large cyanogen bromide
RT   fragments of sei whale (Balaenoptera borealis) somatotropin.";
RL   Bioorg. Khim. 4:1589-1599(1978).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
DR   PIR; PN0140; PN0140.
DR   PIR; JN0387; JN0387.
DR   HSSP; P01241; 1AXI.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00466; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary.
FT   DISULFID     52    164       BY SIMILARITY.
FT   DISULFID    189    198       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24815 MW;  69FB8F6B814A7516 CRC64;

  Query Match             65.4%;  Score 407.5;  DB 1;  Length 216;
  Best Local Similarity   67.0%;  Pred. No. 1.1e-25;

  Matches   61;  Conservative   14;  Mismatches   15;  Indels    1;  Gaps    1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPTP 61
           ||  ||| ||  |||||  ||||| |||:|||  |||:  |:|  |||  ||  ||:  |||:||
Db      27 FPAMPLSNLFANAVLRAQHLAQLAADTFKEFERTYIPEGQRYSIQNAQAAFCFSETIPAP 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |:|:|| |||:|||| |||||||||| |||
Db      87 PGKEEAQQRSDMELLRFSLLLIQSWLGPVQ 116


RESULT 12
SOMA_MOUSE
ID   SOMA_MOUSE     STANDARD;      PRT;   216 AA.
AC   P06880;
DT   01-JAN 1988 (Rel. 06, Created)
DT   01-JAN 1988 (Rel. 06, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gurevich V.S., Tsukanova L.I., Kuzina I.I.;
RA   Linzer D.I.H., Talamantes F.;
RT   "Nucleotide sequence of mouse prolactin and growth hormone mRNAs and
RT   expression of these mRNAs during pregnancy.";
RL   J. Biol. Chem. 260:9574-9579(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ZTED; TISSUE=Liver;
RX   MEDLINE=96191857; PubMed=8647118;
RA   Dar P., Meyer L., Goffet H. D., Ben-Oman S., Pfeffer M.;
RT   "Structure of the growth hormone encoding gene and its promoter in
RT   mice.";
RL   Gene 169:209-213(1996).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AA...; CAA26650.1; -.
DR   EMBL; ...; CAA86658.1; -.
DR   PIR; B23911; SIMS.
DR   HSSP; P01246; 1BSI.
DR   MIM; 601.95787; GH.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00466; SOMATOTROPIN.
DR   ProSite; PS00266; SOMATOTROPIN_1; 1.
DR   ProSite; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26       BY SIMILARITY.
FT   CHAIN        27    216       SOMATOTROPIN.
FT   DISULFID     76    164       BY SIMILARITY.
FT   DISULFID    206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24716 MW;  98666A4AB2CD65FC CRC64;
```

```
 Query Match             64.8%;  Score 534.5;  DB 1;  Length 216;
 Best Local Similarity   64.8%;  Pred. No. 2.2e-56;
 Matches  95;  Conservative  14;  Mismatches  17;  Indels   1;  Gaps   1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           | ||||  ||  ||||||  |||| |||||| |||| |||  ||| | |  ||||||
Db      27 FPAMPLSNLFVNGVRASQHLMQLSLSTFEEFIHAYIPEPDKKVAFTKSIECQPPPTPVN 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |  ||| | |||||||| |||||||| |||
Db      87 PIGKDENAQQKSDMELLRFSLLLIPSWLGPVQ 117


RESULT 13
SOMA_RABIT
ID   SOMA_RABIT      STANDARD;      PRT:   216 AA.
AC   P46407;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN 2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white;
RX   MEDLINE=91211854; PubMed=2008237;
RA   Wallis E.C., Wallis M.;
RT   "Cloning and characterisation of the rabbit growth hormone and
RT   gene.";
RL   Gene 103:253-256(1991).
CC   -!- FUNCTION: Plays an important role in growth control. It is a
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U02029; AAA51047.1; -.
DR   PIR; A01514; SCRB.
DR   HSSP; P01246; 1HST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00837; SOMATOTROPIN.
DR   PROSITE; PS00264; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL       1     26
FT   CHAIN       27    216       SOMATOTROPIN.
FT   DISULFID     78    190
FT   DISULFID    204    214
SQ   SEQUENCE   216 AA;  24154 MW;  F7B8C1D4E1C1A7B9 CRC64;

  Query Match             64.8%;  Score 534.5;  DB 1;  Length 216;
  Best Local Similarity   64.8%;  Pred. No. 2.2e-56;
  Matches  95;  Conservative  14;  Mismatches  17;  Indels   1;  Gaps   1;

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 61
           | ||||  ||  ||||||  |||| |||||| |||| |||  ||| | |  ||||||
Db      27 FPAMPLSSLFANAVLRAQHLHQLAADTYKEFERAYIPEDQKYSFLQNPQTSLCFSESIPT 86

Qy      62 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |  ||| | |||||||| |||||||| |||
Db      87 PIGKDENAQQKSDMELLRFSLLLIPSWLGPVQ 117


RESULT 14
SOMA_RABIT
ID   SOMA_RABIT      STANDARD;      PRT:   216 AA.
AC   P46407;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUN 2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white;
RX   MEDLINE=91211854; PubMed=2008237;
RA   Wallis E.C., Wallis M.;
RT   "Cloning and characterisation of the rabbit growth hormone and
RT   gene.";
RL   Gene 103:253-256(1991).
CC   -!- FUNCTION: Plays an important role in growth control. It is a
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE SOMATOTROPIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z50124; CAA90248.1; -.
DR   HSSP; P01241; 1HST.
DR   InterPro; IPR001400; Somatotropin.
DR   Pfam; PF00103; hormone; 1.
```

```
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00340, SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26       POTENTIAL.
FT   CHAIN        27    216       SOMATOTROPIN.
FT   DISULFID     78    189       BY SIMILARITY.
FT   DISULFID    206    214       BY SIMILARITY.
SQ   SEQUENCE   216 AA;  24114 MW;  9E7167481796F075 CRC64;

  Query Match             64.4%;  Score 302.5;  DB 1;  Length 216;
  Best Local Similarity   64.8%;  Pred. No. 4.3e-25;
  Matches   59;  Conservative   11;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNTQTSLCFSESIPT 61
           ||  ||| || ||:|||  |||:| |||:|| |||: |:|| :|| | :  |||:||
Db      27 PPAMDNRQFFHRSVLKVIHDEEQEMLKKPLCFTDDVKMAPAQKRAAEIKKDIQAIQVAGA 86

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
           |: ::| |||:|:||| |||||||| |||
Db      87 PIGKEDAEQSNMMELSFQSGLEIQLNRLLDG 116


RESULT 15
SOMA_BAT
ID   SOMA_BAT        STANDARD;      PRT;   216 AA.
AC   P01244;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Somatotropin precursor (Growth hormone).
GN   GH1 OR GH.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=82059626; PubMed=6272224;
RA   Page G.S., Smith S., Goodman H.M.;
RT   "DNA sequence of the rat growth hormone gene: location of the 5'
RT   terminus of the growth hormone mRNA and identification of an internal
RT   transposon like element.";
RL   Nucleic Acids Res. 9:2887-2901(1981).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=78071760; PubMed=4910;
RA   Seeburg P.H., Shine J., Martial J.A., Baxter J.D., Goodman H.M.;
RT   "Nucleotide sequence and amplification in bacteria of structural gene
RT   for rat growth hormone.";
RL   Nature 270:486-494(1977).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=82066155; PubMed=6946433;
RA   Barta A., Richards R.I., Baxter J.D., Shine J.;
RT   "Primary structure and evolution of rat growth hormone gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:4867-4871(1981).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=96056684; PubMed=8521139;
RA   Roth W.M., Weigent D.A.;
RT   "Cloning and nucleotide sequencing of rat lymphocyte growth hormone
RT   cDNA.";
RL   Neuroimmunomodulation 2:108-114(1995).
CC   -!- FUNCTION: Plays an important role in growth control. Its major
CC       role in stimulating body growth is to stimulate the liver and
CC       other tissues to secrete IGF-1. It stimulates both the
CC       differentiation and proliferation of myoblasts. It also stimulates
CC       amino acid uptake and protein synthesis in muscle and other
CC       tissues.
CC   -!- SUBCELLULAR LOCATION: Secreted.
```

```
CC   -!- SIMILARITY: Belongs to the somatotropin/prolactin family.
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------- ---------- ---------- ---------- ----------
DR   EMBL; V01267; CAA24547.1; -.
DR   EMBL; V01268; CAA24548.1; -.
DR   EMBL; V01270; CAA24549.1; -.
DR   EMBL; M26779; AAB04025.1; -.
DR   PIR; A27511; A27511.
DR   HSSP; P01246; 1BSF.
DR   InterPro; IPR001400; Somato_ropin.
DR   Pfam; PF00103; hormone; 1.
DR   PRINTS; PR00836; SOMATOTROPIN.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
DR   PROSITE; PS00340; SOMATOTROPIN_2; 1.
KW   Hormone; Pituitary; Signal.
FT   SIGNAL        1     26
FT   CHAIN        27    216       SOMATOTROPIN.
FT   DISULFID     78    189       BY SIMILARITY.
FT   DISULFID    206    214       BY SIMILARITY.
FT   CONFLICT     27     27       F -> L (IN REF. 2 AND 4).
SQ   SEQUENCE   216 AA;  24656 MW;  [illegible] CRC64;

  Query Match             64.4%;  Score 302.5;  DB 1;  Length 216;
  Best Local Similarity   64.8%;  Pred. No. 4.3e-25;
  Matches   59;  Conservative   11;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       2 PFTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNTQTSLCFSESIPT 61
           ||  ||| || ||:|||  |||:| |||:|| |||: |:|| :|| | :  |||:||
Db      27 PPAMDNRQFFHRSVLKVIHDEEQEMLKKPLCFTDDVKMAPAQKRAAEIKKDIQAIQVAGA 86

Qy      62 PSNKEETQQKSNLELLRISLLLIQSWLEPVQ 92
           |: ::| |||:|:||| |||||||| |||
Db      87 PIGKEDAEQSNMMELSFQSGLEIQLNRLLDG 116


Search completed: February  5, 2003, 12:44:57
Job time : 5.23473 secs
```

Run on: February 5, 2003, 12:42:21 ; Search time 0.0072 Seconds
(without alignments)
7.712* Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 47
Sequence: 1 MEFTIPLSFEFDNAMLMAHE.....NLEELPCSELIISWLEEVQ 92

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Query Match: 100.0% Score 47 / 92 Length

Number of hits satisfying chosen parameters:

Minimum DB seq length:
Maximum DB seq length:

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_synthetic:*
13: sp_unclassified:*
14: sp_virus:*
15: sp_bacteriaj:*
16: sp_archaej:*

Pred. No. is the number of results predicted by chance to
score higher than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	47	100.0	217	4	gi|WNE|	gi|WNE| Arabis alpina
2	42.5	90.4	248	4	gi|14|	gi|14| Homo sapien
3	40	85.1	217	4	gi||	gi|| Arabis hirs
4	40	85.1	217	4	gi||	gi|| Arabis hirs
5	40	85.1	217	4	gi|14407	gi|14407 Homo sapien
6	38.5	81.9	217	4	gi|WNE|	gi|WNE| Arabis alpina
7	37.5	79.8	222	4	gi|14|	gi|14| Homo sapien
8	37.5	79.8	222	12		
9	37.5	79.8	230	14	gi||	gi|| Human herp
10	37.5	79.8	234	4	gi||	gi|| Arabis hirs
11	37.5	79.8	234	4	gi||	gi|| Arabis hirs
12	37.5	79.8	234	14	gi||	gi|| Human herp
13	37.5	79.8	234	14	gi||	gi|| Human herp
14	37.5	79.8	234	14	gi||	gi|| Human herp
15	37.5	79.8	234	14	gi||	gi|| Human herp
16	37.5	79.8	234	14	gi||	gi|| Human herp
17	37.5	79.8	234	14	gi||	gi|| Human herp
18	37.5	79.8	234	14	gi||	gi|| Human herp
19	37.5	79.8	234	14	gi||	gi|| Human herp
20	37.5	79.8	234	14	gi||	gi|| Human herp
21	37.5	79.8	234	14	gi||	gi|| Human herp
22	37.5	79.8	234	14	gi||	gi|| Human herp
23	37.5	79.8	234	14	gi||	gi|| Human herp
24	37.5	79.8	234	14	gi||	gi|| Human herp
25	37.5	79.8	234	14	gi||	gi|| Human herp
26	37.5	79.8	234	14	gi||	gi|| Human herp
27	37.5	79.8	234	14	gi||	gi|| Human herp
28	37.5	79.8	234	14	gi||	gi|| Human herp
29	37.5	79.8	234	14	gi||	gi|| Human herp
30	37.5	79.8	234	14	gi||	gi|| Human herp
31	37.5	79.8	234	14	gi||	gi|| Human herp
32	37.5	79.8	234	14	gi||	gi|| Human herp
33	37.5	79.8	234	14	gi||	gi|| Human herp
34	37.5	79.8	234	14	gi||	gi|| Human herp
35	37.5	79.8	234	14	gi||	gi|| Human herp
36	37.5	79.8	234	14	gi||	gi|| Human herp
37	37.5	79.8	234	14	gi||	gi|| Human herp
38	37.5	79.8	234	14	gi||	gi|| Human herp
39	37.5	79.8	234	14	gi||	gi|| Human herp
40	37.5	79.8	234	14	gi||	gi|| Human herp
41	37.5	79.8	234	14	gi||	gi|| Human herp
42	37.5	79.8	234	14	gi||	gi|| Human herp
43	37.5	79.8	234	14	gi||	gi|| Human herp
44	37.5	79.8	234	14	gi||	gi|| Human herp
45	37.5	79.8	234	14	gi||	gi|| Human herp

ALIGNMENTS

RESULT 1
gi|WNE|
ID gi|WNE| standard; DNA; 217 BP.
AC gi|WNE|;
DT 01-MAR-2002 (Rel. 1, Created)
DT 01-MAR-2002 (Rel. 1, Last sequence update)
DT 01-JAN-2002 (Rel. 1, Last annotation update)
DE Arabis alpina
KW .
OS Arabis alpina (Alpine rockcress).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OC Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA Roles A.J., Ingvarsson P.K., Gaut B.S., Purugganan M.D.;
RT "Independent duplication of the gi|WNE| homeodomain in Poaceae
RT and Arabis petraea.";
RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AF[illegible] AAG[illegible].
RL [illegible]; [illegible] [illegible].
DR PFAM; PF[illegible]; [illegible].
DR PROSITE; PS[illegible]; [illegible].
DR INTERPRO; IPR[illegible]; [illegible].
DR PIR; PI[illegible]; [illegible].
DR PIR; PI[illegible]; [illegible].
SQ SEQUENCE 217 AA; [illegible] MW; [illegible] CRC64;

Query Match: [illegible]% Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]%; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

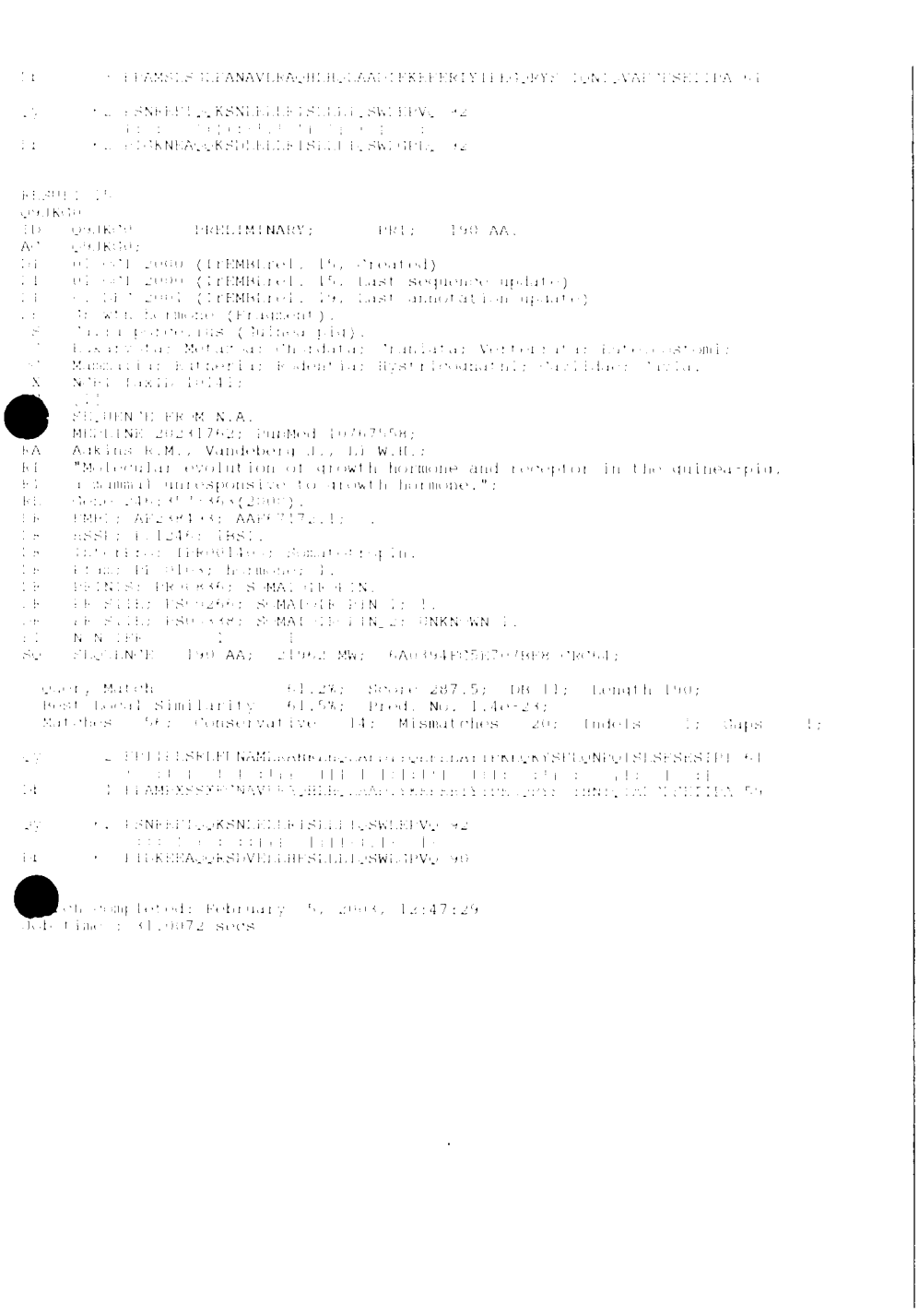
Qy [illegible]
Db [illegible]

Qy [illegible]

Db [illegible]

```
RESULT 3
...
ID   ...      PRELIMINARY;      PRT:   ... AA.
AC   ...;
DT   01-NOV-1996 (IrEMBLrel. 01, Created)
DT   01-NOV-1996 (IrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE   Placental growth hormone isoform hGH V8 precursor.
GN   HGH V.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=98037737; PubMed=9792964;
RA   Barbour L.A., Shao J., Qiao L., Pierce W.M., Jr., Duckett R.J.,
RA   Cardiff R.D., Anthony R.V., Jaeger L.A., Clark R.,
RA   Caruso-Nicoletti M., Catalano P.M., Hankins G.D.;
RT   "Human placental growth hormone-variant is expressed in human
RT   breast tumors.";
RL   J. Clin. Endocrinol. Metab. 83:3828-3835(1998).
CC   -!- MASS_SPECTROMETRY: MW=27101; METHOD=Unknown;
CC   -!- TISSUE SPECIFICITY: PLACENTA.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   245 AA;  27101 MW;  14CC7F9C02191C08 CRC64;

  Query Match             89.8%;  Score 422;  DB 4;  Length 245;
  Best Local Similarity   92.4%;  Pred. No. 4e-38;
  Matches  214;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 MATPSKTSCFLMILVLSFFHVASSAIPIGPVQRLHDSGIFKDNLTKVLLSNLLKNYRKFIP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MATPSKTSCFLMILVLSFFHVASSAIPIGPVQRLHDSGIFKDNLTKVLLSNLLKNYRKFIP 60

Qy      61 SNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |||||||||||||||||||||||||||||||
Db      61 SNREETQQKSNLELLRISLLLIQSWLEPVQ 92


RESULT 4
...
ID   ...      PRELIMINARY;      PRT:   217 AA.
AC   ...;
DT   01-NOV-1996 (IrEMBLrel. 01, Created)
DT   01-NOV-1996 (IrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (IrEMBLrel. 19, Last annotation update)
DE   Somatotropin 5 precursor (Growth hormone 5).
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RX   MEDLINE=94009824; PubMed=8047117;
RA   Talamantes F., Barnhart M., Fisher J.M., Power R.F.,
RT   "Placental growth hormone/deficient somatomammotropin variant:
RT   comparison of expression and secretion of this differentially expressed human
RT   protein in the rhesus monkey placenta.";
RL   Endocrinology 134:1741-1750(1994).
CC   -!- MASS_SPECTROMETRY: MW=22554; METHOD=Unknown;
CC   -!- TISSUE SPECIFICITY: PLACENTA.
DR   PROSITE; PS00266; SOMATOTROPIN_1; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   217 AA;  24471 MW;  F1EC0AEC08A1C1BC CRC64;

  Query Match             89.8%;  Score 422;  DB 4;  Length 217;
  Best Local Similarity   92.4%;  Pred. No. 3.5e-38;
  Matches  214;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 MATPSKTSCFLMILVLSFFHVASSAIPIGPVQRLHDSGIFKDNLTKVLLSNLLKNYRKFIP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MATPSKTSCFLMILVLSFFHVASSAIPIGPVQRLHDSGIFKDNLTKVLLSNLLKNYRKFIP 60

Qy      61 SNREETQQKSNLELLRISLLLIQSWLEPVQ 92
           |||||||||||||||||||||||||||||||
Db      61 SNREETQQKSNLELLRISLLLIQSWLEPVQ 92


RESULT 5
...
ID   ...      PRELIMINARY;      PRT:   ... AA.
AC   ...;
```

```
ID   [illegible]      PRELIMINARY;      PRT;   202 AA.
AC   [illegible];
DT   31-JAN-1998 (Rel. 05, Created)
DT   01-JAN-1998 (Rel. 05, Last sequence update)
DT   15-NOV-2002 (Rel. 21, Last annotation update)
DE   [illegible] growth hormone 2 [illegible].
ON   Bos sp.
OS   Bos sp. (cattle).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Laurasiatheria; Artiodactyla; Ruminantia; Pecora;
OC   Bovidae; Bos.
OX   NCBI_TaxID=9906;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Carlsson L.M.S., Carlsson B.;
RT   "Cloning of two novel growth hormone transcripts expressed in human
RT   placenta.";
RL   J. Clin. Endocrinol. Metab. 83:2978-2985(1998).
DR   EMBL; AF[illegible]; AAB71828.1; -.
DR   HSSP; P01241; 1A22.
DR   InterPro; IPR001400; Somatotrophin.
DR   Pfam; PF00103; hormone_1.
DR   PRINTS; PR00837; S-MAMMOTROPIN.
DR   PROSITE; PS00266; S-MAMMOTROPIN_1; 1.
DR   PROSITE; PS00267; S-MAMMOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
SQ   SEQUENCE   202 AA;  [illegible] MW;  [illegible];

  Query Match          71.8%;  Score 239.5;  DB 4;  Length 202;
  Best Local Similarity   76.9%;  Pred. No. 6.7e-29;
  Matches   70;  Conservative    5;  Mismatches   [illegible];  Indels   [illegible];  Gaps    1;

Qy       1 MFTIISLFLNAMLAHEAHQAISNTYRQIFGDAYQKFIQNPQTSLSFSESIPT 51
           | | | || |||||| ||||||||||||||||||||||||||||||||||||
Db      [illegible] MFTIISLFLNAMLAHEAHQAISNTYRQIFGDAYQKFIQNPQTSLSFSESIPT 71

Qy      52 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 82
           |||||||||||||||||||||||||||||||
Db      72 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 102


RESULT 2
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   216 AA.
AC   [illegible];
DT   01-AUG-1998 (Rel. 07, Created)
DT   01-AUG-1998 (Rel. 07, Last sequence update)
DT   15-NOV-2002 (Rel. 19, Last annotation update)
DE   Growth hormone precursor.
OS   Spalax ehrenbergi ehrenbergi (Ehrenberg's mole rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciuromorpha;
OC   Muroidea; Spalacidae.
OX   NCBI_TaxID=10161;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99243164; PubMed=10224117;
RA   Hoagland A., Nevo E., Wallis M.;
RT   "Cloning and characterisation of the gene encoding mole rat (Spalax
RT   ehrenbergi) growth hormone.";
RL   J. Mol. Endocrinol. 22:29-36(1999).
DR   EMBL; AF056829; AAA06716.1; -.
DR   HSSP; P01241; 1BSI.
DR   InterPro; IPR001400; Somatotrophin.
DR   Pfam; PF00103; hormone_1.
DR   PRINTS; PR00837; S-MAMMOTROPIN.
DR   PROSITE; PS00266; S-MAMMOTROPIN_1; 1.
DR   PROSITE; PS00267; S-MAMMOTROPIN_2; 1.
KW   Signal.
FT   SIGNAL       1     26       POTENTIAL.
FT   CHAIN       27    216       GROWTH HORMONE.
SQ   SEQUENCE   216 AA;  [illegible] MW;  [illegible];

  Query Match          [illegible];  Score 208.5;  DB [illegible];  Length 216;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 62
           | | | || |||||| |||||||||||||||||||||||||||||||||||||
Db      [illegible] FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 77

Qy      63 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 93
           |||||||||||||||||||||||||||||||
Db      78 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 108


RESULT 3
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   216 AA.
AC   [illegible];
DT   01-MAY-2000 (Rel. [illegible], Created)
DT   01-MAY-2000 (Rel. [illegible], Last sequence update)
DT   01-JUN-2001 (Rel. [illegible], Last annotation update)
DE   Growth hormone.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciuromorpha;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nguyen T.N.K., Liebhaber S.A.;
RT   "Mouse Growth Hormone Locus: Nucleotide Sequence and Hypersensitive
RT   Analyses.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U34362; AAC00808.1; -.
DR   HSSP; P01241; 1BSI.
DR   InterPro; IPR001400; Somatotrophin.
DR   Pfam; PF00103; hormone_1.
DR   PRINTS; PR00837; S-MAMMOTROPIN.
DR   PROSITE; PS00266; S-MAMMOTROPIN_1; 1.
DR   PROSITE; PS00267; S-MAMMOTROPIN_2; 1.
SQ   SEQUENCE   216 AA;  [illegible] MW;  [illegible];

  Query Match          [illegible];  Score 208.5;  DB [illegible];  Length 216;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy       2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLCFSESIPT 62
           | | | || |||||| |||||||||||||||||||||||||||||||||||||
Db      [illegible] FPAMPLSNLFANAVLRAQHLHQLAADTYKEFERAYIPEGQRYSIQNAQAAFCFSETIPA 77

Qy      63 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 93
           |||||||||||||||||||||||||||||||
Db      78 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 108


RESULT 4
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   01-FEB-1997 (Rel. [illegible], Created)
DT   01-FEB-1997 (Rel. [illegible], Last sequence update)
DT   01-DEC-2001 (Rel. [illegible], Last annotation update)
DE   Placental growth hormone (Fragment).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible];
RX   MEDLINE=97[illegible]; PubMed=[illegible];
```



Run on: February 5, 2003, 12:42:26 ; Search time 3.5167 Seconds
(without alignments)
205.219 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MEPETLSRETNAMLRAm.....NLPTWvSLIL,SWLtV, 4

Scoring table: BL...SUM62
Xgap 10.0, Ygap 0.5

Searched: 2.2574 seqs, 2.1.2722 residues

Total number of hits satisfying chosen parameters: 2.1.74

Minimum DB seq length: 0
Maximum DB seq length: 2...99

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/7A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles2.pep:*

Query No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result printed.
The E-values obtained by analysis of the database distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	1.5.5	1.2.2	1	US-.....-...	Sequence 1, Appli
2	1.1.5	1.1.1	1	US-.....-...	Sequence 5, Appli
3	1.1.5	1.1.1	1	US-.....-...	Sequence 4, Appli
4	1.1.5	1.1.1	1	US-.....-...	Sequence 4, Appli
5	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
6	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
7	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
8	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
9	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
10	1.1.5	1.1.1	1	US-.....-...	Sequence 26, Appl
11	1.1.5	1.1.1	1	US-.....-...	Sequence 71, App.
12	1.1.5	1.1.1	1	US-.....-...	Sequence 73, App.
13	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
14	1.1.5	1.1.1	1	US-.....-...	Sequence 4, Appli
15	1.1.5	1.1.1	1	US-.....-...	Sequence 51, App.
16	1.1.5	1.1.1	1	US-.....-...	Sequence 51, App.
17	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
18	1.1.5	1.1.1	1	US-.....-...	Sequence 18, Appl
19	1.1.5	1.1.1	1	US-.....-...	Sequence 16, Appl
20	1.1.5	1.1.1	1	US-.....-...	Sequence 20, Appl
21	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
22	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
23	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
24	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
25	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
26	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
27	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
28	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
29	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
30	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
31	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
32	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
33	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
34	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
35	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
36	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
37	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
38	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
39	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
40	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
41	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
42	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
43	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
44	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli
45	1.1.5	1.1.1	1	US-.....-...	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-.....-...
; Sequence 1, Application US/...
; Patent No. .....
; GENERAL INFORMATION:
; APPLICANT: ...
; APPLICANT: ...
; APPLICANT: ...
; TITLE OF INVENTION: DNA ... Expression ...
; NUMBER OF SEQUENCES: ...
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ...
; STREET: ...
; CITY: ...
; STATE: ...
; COUNTRY: USA
; ZIP: ...
; COMPUTER READABLE FORM:
; MEDIUM TYPE: ...
; COMPUTER: ...
; OPERATING SYSTEM: ...
; SOFTWARE: ...
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/...
; FILING DATE: ...
; CLASSIFICATION: ...
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ...
; FILING DATE: ...
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ...
; FILING DATE: ...
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: ...
; FILING DATE: ...
; ATTORNEY/AGENT INFORMATION:
; NAME: ...
; REGISTRATION NUMBER: ...
; REFERENCE/DOCKET NUMBER: ...
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: ...
; TELEFAX: ...
; TELEX: ...
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: ... amino acids
; TYPE: amino acid

; Type: Linear
US-08-423-100-2

Query Match 98.9%; Score 465; DB 1; Length 192;
Best Local Similarity 98.9%; Pred. No. 1.1e-51;
Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAYYTSLSKLGTLNAMLRARHLMLDATTYQSFEEAYTPEAKKEIRKLKYFSLSFSESIF 60
Db 1 MAYYTSLSKLGTLNAMLRARHLMLDATTYQSFEEAYTPEAKKEIRKLKYFSLSFSESIF 60

Qy 61 IPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92
Db 61 IPSNREETQQKSNLELLRISLLLIQSWLEPVQ 92

RESULT 2
US-09-284-878-5
; Sequence 5, Application US/09284878
; Patent No. 6342375
; GENERAL INFORMATION:
; APPLICANT: Khachatrian, Martha Guerrero
; APPLICANT: Saldana, Hugo Barrera
; APPLICANT: Salvado, Jose Maria Viader
; TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the
; TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone
; FILE REFERENCE: 1829.0010000
; CURRENT APPLICATION NUMBER: US/09/284,878
; CURRENT FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: 1?/????,???
; PRIOR FILING DATE: 1997-10-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-284-878-5

Query Match 97.9%; Score 460; DB 4; Length 191;
Best Local Similarity 98.9%; Pred. No. 4.6e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 61
Db 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 61

Qy 62 PSNREETQQKSNLELLHISLLLIQSWLEPVQ 92
Db 62 PSNREETQQKSNLELLHISLLLIQSWLEPVQ 92

RESULT 3
US-08-xxx-621-4
; Sequence 4, Application US/08838621
; Patent No. 5917072
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Lockwood, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins And Other
; TITLE OF INVENTION: Proteins By Modification of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jr. Estelle J. Iacobus
; STREET: [illegible]
; CITY: Stamford
; STATE: Connecticut
; COUNTRY: U.S.A.
; ZIP: 06904-0060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,621
; FILING DATE: 06-FEB-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/766,142
; FILING DATE: 25-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Iacobus, Estelle J.
; REGISTRATION NUMBER: 31,145
; REFERENCE/DOCKET NUMBER: 31,278-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-321-2756
; TELEFAX: 203 321 2971
; TELEX: 203 710-474-4059
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 194 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-838-621-4

Query Match 97.9%; Score 460; DB 2; Length 194;
Best Local Similarity 98.9%; Pred. No. 4.7e-51;
Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 61
Db 4 FPTIPLSRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKEQKYSFLQNPQTSLSFSESIP 63

Qy 62 PSNREETQQKSNLELLHISLLLIQSWLEPVQ 92
Db 64 PSNREETQQKSNLELLHISLLLIQSWLEPVQ 94

RESULT 4
US-08-459-906-4
; Sequence 4, Application US/08459906
; Patent No. 6010999
; GENERAL INFORMATION:
; APPLICANT: Daley, Michael J.
; APPLICANT: Lockwood, Brian L.
; APPLICANT: Cady, Susan M.
; APPLICANT: Shieh, Hong-Ming
; APPLICANT: Bohlen, Peter
; APPLICANT: Seddon, Andrew P.
; TITLE OF INVENTION: Stabilization of Somatotropins and other
; TITLE OF INVENTION: Proteins by Modification of Cysteine Residues
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,906
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Darryl L.

```
; Sequence 1, Application US/08/[illegible]
; Patent No. 5,[illegible]
; GENERAL INFORMATION:
;   APPLICANT: Newmark, Christopher B.
;   APPLICANT: Balbas, Philippe A.
;   APPLICANT: No, Gilbert Chinhon, Kan, D.
;   APPLICANT: Clark, Samuel A.
;   APPLICANT: Ibhopen, Amiee E.
;   APPLICANT: Quiados, Christian
;   APPLICANT: Krause, Fred
;   TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;   TITLE OF INVENTION: SECRETORY CELL LINES
;   NUMBER OF SEQUENCES: 56
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arter [illegible], Wells & Banks
;     STREET: P.O. Box 1443
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: USA
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/[illegible]
;     FILING DATE: 05-DEC-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/[illegible]
;     FILING DATE: 06-DEC-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Cross, Edmund D.
;     REGISTRATION NUMBER: 34,[illegible]
;     REFERENCE/DOCKET NUMBER: DX-[illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-854-9196
;     TELEFAX: 415-495-1200
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 217 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 42..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 94..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: [illegible]..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 192..217
;     OTHER INFORMATION: /note= "Site of tissue plasminogen activator
;     OTHER INFORMATION: /deposited to [illegible]."
US-08/[illegible]-1

  Query Match          97.9%;  Score 460;  DB 4;  Length 217;
  Best Local Similarity  98.9%;  Pred. No. 5.5e-51;
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy      1 [illegible]
          [illegible]
Db      1 [illegible]

RESULT 2
US/08/[illegible]
; Sequence 1, Application US/08/[illegible]
; Patent No. 6225416
; GENERAL INFORMATION:
;   APPLICANT: Allsopp, Scott W.
;   APPLICANT: Bowdy, J. Fernando
;   APPLICANT: Kostelnik, Robert A.
;   TITLE OF INVENTION: MULTIFUNCTIONAL VALENCY [illegible] HAMELIAN DOMAINS
;   NUMBER OF SEQUENCES: 11
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: DNAX Research Institute
;     STREET: 901 California Avenue
;     CITY: Palo Alto
;     STATE: California
;     COUNTRY: USA
;     ZIP: 94304-1104
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/[illegible]
;     FILING DATE: 05-DEC-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/[illegible]
;     FILING DATE: 06-DEC-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Cross, Edmund D.
;     REGISTRATION NUMBER: 34,[illegible]
;     REFERENCE/DOCKET NUMBER: DX-[illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-854-9196
;     TELEFAX: 415-495-1200
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 217 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 42..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 94..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: [illegible]..[illegible]
;   FEATURE:
;     NAME/KEY: Peptide
;     LOCATION: 192..217
;     OTHER INFORMATION: /note= "Site of tissue plasminogen activator
;     OTHER INFORMATION: /deposited to [illegible]."
US-08/[illegible]-1

  Query Match          97.9%;  Score 460;  DB 4;  Length 217;
  Best Local Similarity  98.9%;  Pred. No. 5.5e-51;
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy      1 [illegible]
          [illegible]
Db      1 [illegible]

RESULT 3
US/09/[illegible]-1
; Sequence 1, Application US/[illegible]
; Patent No. 6342375
; GENERAL INFORMATION:
;   APPLICANT: Chapman, Martha Dorris
```

```
RESULT 12
US-08-241-046-2
; Sequence 2, Application US/08241046
; Patent No. 5611792
; GENERAL INFORMATION:
;   APPLICANT: Howard, Christopher B.
;   APPLICANT: Balloun, Philippe A.
;   APPLICANT: Nau, Fritz/Zimmermann, Karl H.
;   APPLICANT: Clark, Samuel A.
;   APPLICANT: Thiriart, Alice E.
;   APPLICANT: Quinder, Christian
;   APPLICANT: Krause, Fred
;   APPLICANT: McCarry, Dennis
;   TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;   TITLE OF INVENTION: SECRETORY CELL LINES
;   NUMBER OF SEQUENCES: 79
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: USA
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/241,046
;     FILING DATE: Concurrently Herewith
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/028,427
;     FILING DATE: 15-OCT-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/009,616
;     FILING DATE: 19-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weinhardt, Steven L.
;     REGISTRATION NUMBER: 37,942
;     REFERENCE/DOCKET NUMBER: 0150.5411
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 512/418-3000
;     TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 96 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
US-08-241-046-2

  Query Match          97.9%;  Score 1954;  DB 4;  Length 191;
  Best Local Similarity  98.9%;  Pred. No. 1.1e-194;
  Matches   94;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 EPTETLSKELTDLLRADIYSLETEEAYIPKEAKKVSFLQNPQTSLCFSESIPT 54
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 EPTETLSKELTDLLRADIYSLETEEAYIPKEAKKVSFLQNPQTSLCFSESIPT 54

Qy      55 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 85
           |||||||||||||||||||||||||||||||
Db      55 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 85


RESULT 13
US-08-241-046-3
; Sequence 3, Application US/08241046
; Patent No. 5611792
; GENERAL INFORMATION:
;   APPLICANT: Howard, Christopher B.
;   APPLICANT: Balloun, Philippe A.
;   APPLICANT: Nau, Fritz/Zimmermann, Karl H.
;   APPLICANT: Clark, Samuel A.
;   APPLICANT: Thiriart, Alice E.
;   APPLICANT: Quinder, Christian
;   APPLICANT: Krause, Fred
;   APPLICANT: McCarry, Dennis
;   TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM
;   TITLE OF INVENTION: SECRETORY CELL LINES
;   NUMBER OF SEQUENCES: 79
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Arnold, White & Durkee
;     STREET: P.O. Box 4433
;     CITY: Houston
;     STATE: Texas
;     COUNTRY: USA
;     ZIP: 77210
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/241,046
;     FILING DATE: Concurrently Herewith
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/028,427
;     FILING DATE: 15-OCT-1996
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 60/009,616
;     FILING DATE: 19-JAN-1996
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weinhardt, Steven L.
;     REGISTRATION NUMBER: 37,942
;     REFERENCE/DOCKET NUMBER: 0150.5411
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 512/418-3000
;     TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 96 amino acids
;     TYPE: amino acid
;     STRANDEDNESS:
;     TOPOLOGY: linear
US-08-241-046-3

  Query Match          97.9%;  Score 1954;  DB 4;  Length 191;
  Best Local Similarity  98.9%;  Pred. No. 1.1e-194;
  Matches   94;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 EPTETLSKELTDLLRADIYSLETEEAYIPKEAKKVSFLQNPQTSLCFSESIPT 54
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 EPTETLSKELTDLLRADIYSLETEEAYIPKEAKKVSFLQNPQTSLCFSESIPT 54

Qy      55 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 85
           |||||||||||||||||||||||||||||||
Db      55 PSNREETQQKSNLELLRISLLLIQSWLEPVQ 85


RESULT 14
US-08-167-756C-4
; Sequence 4, Application US/08167756C
; Patent No. 5997709
; GENERAL INFORMATION:
;   APPLICANT: BOSEN, ET AL.
;   TITLE OF INVENTION: Human Growth Hormone
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: GENETECH, INC.
;     STREET: 460 POINT SAN BRUNO BLVD
;     CITY: SOUTH SAN FRANCISCO
;     STATE: CALIFORNIA
;     COUNTRY: USA
;     ZIP: 94080
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 INCH DISKETTE
;     COMPUTER: IBM PC/AT
;     OPERATING SYSTEM: MS-DOS
;     SOFTWARE: WORD PERFECT 5.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/167,756C
;     FILING DATE: December 17, 1993
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: FERRARO, GREGORY D.
;     REGISTRATION NUMBER: 36,134
;     REFERENCE/DOCKET NUMBER: 227/PCT/US
;     TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 201 994 1700
;     TELEFAX: 201 994 1744
; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 217 AMINO ACIDS
;     TYPE: AMINO ACID
;     STRANDEDNESS:
;     TOPOLOGY: LINEAR
;   MOLECULE TYPE: PEPTIDE
US-08-167-756C-4
```

GMprotein protein search, using sw model

Run on: February 5, 2003, 12:44:31 ; Search time 0.24828 Seconds
(without alignments)
247.409 Million cell updates/sec

Title: US-09-423-100-2
Perfect score: 470
Sequence: 1 MEPEIPLSKLEDNAMEIARE.....NLELLKSLLLLQWEIVL 92

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 129565 seqs, 22169297 residues

Total number of hits satisfying chosen parameters: 129565

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubaa/US09B_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubaa/US09D_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubaa/US61_PUBCOMB.pep:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	470	100.0	92	9	US-10-054-873-2	Sequence 2, Appli
2	470	100.0	152	9	US-10-054-873-7	Sequence 7, Appli
3	468	99.6	217	10	US-10-039-929-918-7	Sequence 9, Appli
4	460	97.9	245	10	US-09-290-022-46	Sequence 66, Appl
5	451	96.0	217	10	US-09-854-788-2	Sequence 2, Appli
6	450	95.7	217	9	US-09-804-409A-16	Sequence 16, Appl
7	447	95.1	217	10	US-09-854-788-4	Sequence 4, Appli
8	396	84.3	217	10	US-09-854-788-3	Sequence 3, Appli
9	266	56.6	49	9	US-10-054-873-1	Sequence 1, Appli
10	260	55.3	107	9	US-10-054-873-6	Sequence 6, Appli
11	194.5	41.4	66	10	US-09-887-509A-2	Sequence 2, Appli
12	112	23.8	54	10	US-09-876-428-14	Sequence 14, Appl
13	105	22.3	198	9	US-10-146-293-52	Sequence 52, App.
14	104	22.1	199	9	US-10-146-293-16	Sequence 16, Appl
15	104	22.1	199	9	US-10-146-293-26	Sequence 26, Appl
16	126	26.8	199	9	US-10-146-293-27	Sequence 27, App.
17	124	26.4	199	9	US-10-146-293-31	Sequence 31, Appl
18	124	26.4	199	9	US-10-146-293-22	Sequence 22, Appl
19	122	26.0	199	9	US-10-146-293-21	Sequence 21, Appl
20	121.5	25.9	52	10	US-09-[illegible]	Sequence 23, Appl
21	121.5	25.9	52	10	US-09-[illegible]	Sequence 27, Appl
22	119	25.3	199	9	US-10-146-293-53	Sequence 53, Appl
23	118	25.1	199	9	US-10-146-293-29	Sequence 29, Appl
24	118	25.1	199	9	US-10-146-293-17	Sequence 17, Appl
25	117	24.9	197	9	US-10-146-293-15	Sequence 15, Appl
26	116	24.7	197	9	US-10-146-293-14	Sequence 14, Appl
27	116	24.7	199	9	US-10-146-293-2	Sequence 2, Appl
28	116	24.7	199	9	US-10-146-293-12	Sequence 12, Appl
29	116	24.7	199	9	US-10-146-293-25	Sequence 25, Appl
30	116	24.7	227	9	US-10-146-293-5	Sequence 5, Appl
31	116	24.7	227	9	US-10-146-293-4	Sequence 4, Appl
32	116	24.7	227	10	US-09-[illegible]	Sequence 1, Appl
33	115	24.5	199	9	US-10-146-293-9	Sequence 9, Appl
34	115	24.5	199	9	US-10-146-293-7	Sequence 7, Appl
35	114	24.3	54	10	US-09-[illegible]	Sequence 1, Appl
36	114	24.3	54	10	US-09-[illegible]	Sequence 2, Appl
37	114	24.3	199	9	US-10-146-293-12	Sequence 12, Appl
38	112	23.8	199	9	US-10-146-293-10	Sequence 10, Appl
39	111	23.6	198	9	US-10-146-293-55	Sequence 55, Appl
40	108	22.9	28	10	US-09-[illegible]	Sequence 4, Appl
41	106	22.6	28	10	US-09-[illegible]	Sequence 7, Appl
42	106	22.6	54	10	US-09-[illegible]	Sequence 3, Appl
43	101	21.5	198	9	US-10-146-293-54	Sequence 54, Appl
44	101	21.5	197	9	US-10-146-293-18	Sequence 18, Appl
45	99	21.1	199	9	US-10-146-293-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-10-054-873-2
; Query Match 100.0%
; Perfect match 100.0%
; Length 92 ; Initial
; GENERAL INFORMATION:
; Applicant: [illegible]
; TITLE OF INVENTION: Thiocarboxylic Acid Containing and Thiazolidine Derivatives Useful as Chaperone-like Sequences
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/[illegible]
; FILING DATE: 26-MAR-1997
; APPLICATION NUMBER: US 08/[illegible]
; FILING DATE: 11-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Keller, Frank A.
; REGISTRATION NUMBER: 41,494
; REFERENCE/DOCKET NUMBER: [illegible]
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 92 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

; Type: amino acid
; STRANDEDNESS: single
; Topology: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; NAME: [illegible]
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
[illegible]

Query Match 84.3%; Score 396; DB 10; Length 217;
Best Local Similarity 84.6%; Pred. No. 1.2e-38;
Matches 77; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 2
[illegible]
; Sequence 5, Application PCT/US2002016471A1
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Sun, Zhong F.
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Frank J
; REGISTRATION NUMBER: 44,964
; REFERENCE/DOCKET NUMBER: 020267-000500US
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 217 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US 10/054,873-6

Query Match 55.9%; Score 262; DB 9; Length 217;
Best Local Similarity 100.0%; Pred. No. 2.8e-24;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 12
US-09-887-569A-2
; Sequence 2, Application US/09887569A
; Patent No. US20020100073A1
; GENERAL INFORMATION:
; APPLICANT: McCuaig, Murray R.
; APPLICANT: Habibi, Hamid R.
; TITLE OF INVENTION: Expression of Somatotropin in Food Seeds
; FILE REFERENCE: 9369.183
; CURRENT APPLICATION NUMBER: US/09/887,569A
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/214,043
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 60
; TYPE: PRT

```
                 : :  ||| |:  :| :| :: : : ||:| |    |   |:          | |: :| :
Db      13 VSLGELFDRAVELSHYIHFLSERIPNNPEKEV  -AQCRKDFTRAVNGCHPN-SLIPPIE 68

Qy      65 REETQQKSNLELLRISLLLIQSWLEPV 91
           :|:|||  : ||| : | :::|| :|:
Db      69 REQIQQTHHEELLNLCLLQVLRGWNDPL 95


Search completed: February 5, 2003, 12:48:47
Job time : 8.2437s secs
```

Run on: February 5, 2003, 12:42:21 ; Search time 45.1498 Seconds
(without alignments)
526.113 Million cell updates/sec

[redacted] US-09-423-100-4

Score: 463

Sequence: 1 EVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 99

Scoring table: BLOSUM62
Gap 10.0, Ext 0.5

Searched: 4 seqs, 1002 res

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 2

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*

1: /SIDS2/gcgdata/gcgb[illegible]
2: [illegible]
3: [illegible]
4: [illegible]

The number to the left of each hit is the hit number. The score is the number of matches found between the query sequence and the hit sequence. The percent score is the score divided by the number of residues in the query sequence. The percent match is the score divided by the number of residues in the shorter of the two sequences. The last column is the description of the hit sequence.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	463	100.0	99	1	AAB19629	Sequence of human
2	463	100.0	99	2	AAE14613	Human insulin. Hu
3	463	100.0	99	2	AAY12476	Human insulin pro
4	463	100.0	99	21	AAE12613	Human pi insulin
5	463	100.0	99	21	AAM39613	Human pi insulin
6	463	100.0	99	3	AAB12629	Human proinsulin
7	463	100.0	99	5	AAB12613	Sequence of the s2
8	463	100.0	99	6	AAB12613	Sequence 1 the s2
9	463	100.0	99	1	AAE12613	/cds= (1..99) pi Gen
10	463	100.0	99	1	AAE12613	Sequence 1 the s2

ALIGNMENTS

RESULT 1
AAB19629
ID AAB19629 standard protein 99 AA.
XX
AC AAB19629;
XX
DT 27-JUL-1995 (first entry)
XX
DE Sequence 1 from Patent WO9514715.
XX
KW Insulin precursor; diabetes; hypoglycaemic; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region /label= Chain B
FT Peptide /label= Insulin
FT /note= "Insulin peptide"
FT Region /label= Chain A
FT Modified-site /label= 1 NH2
FT /label= "N-terminal Phe"
FT /note= "N-terminal AA residue is pyroglu residue"
FT Disulfide bond /label= Cys-Cys
FT Disulfide bond /label= Cys-Cys
FT Disulfide bond /label= Cys-Cys
FT Modified-site /label= NH2
XX
SQ Sequence 99 AA;
XX

```
RL    ... FEB 2003.
XX
RF    D. FEB 2003;    8,005,034,997.
XX
PR    ... FEB 1999;    9SUS-034997.
PR    ... MAR 1999;    98US-134000.
PR    ... NOV 1999;    99US-0210696.
XX
PA    (GENE) ELI LILLY & CO.
XX
DR    Issue 980.
XX
PT    Insulin precursor produc. from linear S-sulphonate and mercaptan -
PT    in single step without separate oxidn.
XX
PS    Claim 17; Column 41; 89pp; English.
XX
CC    The inventors claim a method for the prepn. of an insulin precursor
CC    in which the A chain and B chain are joined through a connecting
CC    peptide. The connecting peptide joins the A chain at the amino
CC    group of A1 to the B chain at the carboxy, at up of B 30. The
CC    method is prel. for the prepn. of human insulin precursor (see
CC    AAP10020). The SQs of the connecting peptides of a number of species
CC    are given. (see AAP40020, AAP40030 etc).
XX
SQ    Sequence    97 AA;

  Query Match              100.0%;  Score 463;  DB 5;  Length 86;
  Best Local Similarity    100.0%;  Pred. No. 4.9e-45;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
           ||||||||||||||||||||||||||
Db      61 SLQKRGIVEQCCTSICSLYQLENYCN 86


RESULT 2
AAP84064
ID    AAP84064 standard; protein; 86 AA.
XX
AC    AAP84064;
XX
DT    19-JAN-2001  (first entry)
XX
DE    Human insulin precursor, SEQ ID 5.
XX
KW    Insulin precursor; growth hormone; hyperglycaemia; diabetes;
KW    C-peptide; retinopathy; human growth hormone deficiency;
KW    production; yield.
XX
OS    Homo sapiens.
XX
RN    [1]
XX
RD    07-OCT-1999.
XX
RP    1-MAR-1999;  99WO-EP01412.
XX
RP    31-MAR-1998;  98WO-CN000042.
XX
RA    (BOEH) BOEHRINGER MANNHEIM GMBH; CN; CHI.
XX
RT    Gan Z;
XX
RL    WPI; 2000-10639/54.
XX
PT    New chimeric proteins containing human growth hormone fragments, and
PT    particularly for the production of human insulin
XX
PS    Claim 1; Page 22; 24pp; English.
XX
CC    This sequence represents a human insulin precursor comprising insulin
CC    A and B chains separated by a 34 residue peptide sequence. This insulin
CC    precursor can be recombined or chimeric proteins which contain
CC    contains an N terminal fragment of human growth hormone (hGH) and a

RESULT 3
AAY13656
ID    AAY13656 standard; protein; 86 AA.
XX
AC    AAY13656;
XX
DT    19-JAN-2001  (first entry)
XX
DE    Human insulin precursor, SEQ ID 5.
XX
KW    Insulin precursor; growth hormone; hyperglycaemia; diabetes;
KW    C-peptide; retinopathy; human growth hormone deficiency;
KW    production; yield.
XX
OS    Homo sapiens.
XX
RN    [1]
XX
RD    07-OCT-1999.
XX
RP    1-MAR-1999;  99WO-EP01412.
XX
RP    31-MAR-1998;  98WO-CN000042.
XX
RA    (BOEH) BOEHRINGER MANNHEIM GMBH; CN; CHI.
XX
RT    Gan Z;
XX
RL    WPI; 2000-10639/54.
XX
PT    New chimeric proteins containing human growth hormone fragments, and
PT    particularly for the production of human insulin
XX
PS    Claim 1; Page 22; 24pp; English.
XX
CC    This sequence represents a human insulin precursor comprising insulin
CC    A and B chains separated by a 34 residue peptide sequence. This insulin
CC    precursor can be recombined or chimeric proteins which contain
CC    contains an N terminal fragment of human growth hormone (hGH) and a
```

```
CC   cleavable peptide linker (AAY42857). The hGH portion of the chimeric
CC   protein acts as an intramolecular chaperone (IMC) for the insulin
CC   portion, enabling it to fold correctly. The cleavable peptide linker
CC   has a C-terminal Arg residue which enables the hGH portion of the
CC   'hybrid' protein to be removed after folding has taken place. Production
CC   of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC   provide human insulin with correctly linked cysteine bridges with
CC   fewer necessary procedural steps, and hence results in a higher yield
CC   than the cells. The IMC sequences not only protect insulin sequences
CC   from intracellular degradation by a microorganism host, but also promote
CC   the folding of the fused insulin precursor, facilitate the solubility of
CC   the fusion protein and decrease the intermolecular interactions among
CC   the fusion proteins, thus avoiding folding of the fused insulin precursor
CC   at relatively mild fermentation conditions. The procedural steps of
CC   cyanogen bromide cleavage, acid/base sulphitolysis and related
CC   purification steps can thus be eliminated, along with the use of high
CC   concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ   Sequence  ** AA;

  Query Match             Score 4[illegible]; DB 2[illegible]; Length **;
  Best Local Similarity   [illegible]%;  Pred. No. 6.6e-44;
  Matches  **;  Conservative  **;  Mismatches  **;  Indels  **;  Gaps  **;

Qy       1 [illegible]

Db       1 [illegible]

Qy      ** SL_KKRIVEL_[illegible]LENY*N **
           [illegible]
Db      ** SL_KKRIVEL_[illegible]LENY*N **


RESULT 1
AAM42857
ID   AAM42857 standard protein ** AA.
XX
AC   AAM42857;
XX
DT   ** MAR-[illegible] (first entry)
XX
DE   Human proinsulin protein sequence SEQ ID NO:2.
XX
KW   human insulin-like growth factor 1; IGF-1; protein therapy;
KW   diabetic retinal neovascular-like growth factor 1 binding protein; IGFBP-1;
KW   IGFBP-1 stabilizer; hair protective agent; transplantation;
KW   cardiology; anaemia; candidate nephrotropic; dermatological; anti-HIV;
KW   antiviral; hyperglycaemia; obesity; lung disease; glomerulonephritis;
KW   arteriosclerosis; hepatitis; Turner's syndrome; Laron's syndrome;
KW   short stature; increased fat mass; [illegible] disorder;
KW   peripheral neuropathy; multiple sclerosis; muscular dystrophy;
KW   cachexia; stress; trauma; wounding; infection; HIV; skin disorder;
KW   human immunodeficiency virus; diabetes; heart dysfunction;
KW   kidney disorder; whole body growth disorder.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-**
XX
RA   [illegible];
XX
RT   [illegible];
XX
RL   Patent number [illegible];
XX
DR   EMBL; [illegible];
XX
FH   Key             Location/Qualifiers
XX
FT   Peptide         1..**
FT                   /note= "[illegible]"
XX
SQ   Sequence  ** AA;

  Query Match             Score [illegible]; DB 2; Length **;
  Best Local Similarity   [illegible]%;  Pred. No. 6.6e-44;
  Matches  **;  Conservative  **;  Mismatches  **;  Indels  **;  Gaps  **;

Qy       1 [illegible]

Db       1 [illegible]

Qy      ** SL_KKRIVEL_[illegible]LENY*N **
           [illegible]
Db      ** SL_KKRIVEL_[illegible]LENY*N **


RESULT 1
AAM42857
ID   AAM42857 standard protein ** AA
XX
AC   AAM42857;
XX
DT   ** MAR-[illegible] (first entry)
XX
DE   Human proinsulin.
XX
KW   Autoimmune disease; antiviral; anti-parasitic; antidiabetic;
KW   [illegible] with [illegible] [illegible] protein [illegible];
KW   [illegible] [illegible] [illegible] [illegible] [illegible] [illegible];
KW   [illegible] [illegible] [illegible] [illegible] [illegible] [illegible];
XX
OS   Homo sapiens.
XX
RN   [1]
XX
RP   1-**
XX
RA   [illegible] N.V.;
XX
RT   ;
XX
RL   Patent: [illegible]-**-**;
XX
DR   EMBL; MAY [illegible] [illegible];
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..**
FT                   /note= "[illegible]"
XX
SQ   Sequence  ** AA;

  Query Match             Score [illegible]; DB 2; Length **;
  Best Local Similarity   [illegible]%;  Pred. No. 6.6e-44;
  Matches  **;  Conservative  **;  Mismatches  **;  Indels  **;  Gaps  **;

Qy       1 [illegible]

Db       1 [illegible]

Qy      ** SL_KKRIVEL_[illegible]LENY*N **
           [illegible]
Db      ** SL_KKRIVEL_[illegible]LENY*N **


RESULT 1
AAM42857
ID   AAM42857 standard protein ** AA
XX
AC   AAM42857;
XX
DT   ** MAR-[illegible] (first entry)
XX
DE   Human proinsulin.
XX
KW   Autoimmune disease; antiviral; anti-parasitic; antidiabetic;
KW   [illegible] with [illegible] [illegible] protein [illegible];
KW   [illegible] [illegible] [illegible] [illegible] [illegible] [illegible];
XX
OS   Homo sapiens.
XX
RN   [1]
XX
RP   1-**
XX
RA   [illegible] N.V.;
XX
RT   ;
XX
RL   Patent: [illegible]-**-**;
XX
DR   EMBL; MAY [illegible] [illegible];
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..**
FT                   /note= "[illegible]"
XX
SQ   Sequence  ** AA;
```

```
RT   treating cartilage disorders including cartilage damage by injury, a
RT   protective for cartilaginous disorders, by contacting cartilage with
RT   insulin like growth factor analog with altered affinity for IGF binding
RT   proteins.
XX
RS   Sequence 86 AA;

  Query Match             Score 456;  DB 2;  Length 86;
  Best Local Similarity   Score 456;  Pred. No. 1.2e-43;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86


RESULT 2
AAP20096
ID   AAP20096 standard; protein; 87 AA.
XX
AC   AAP20096;
XX
DT   22-JUL-1992  (first entry)
XX
DE   Human preinsulin.
XX
KW   Preinsulin.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-87
RA   Bell G.I.;
RT   ;
RL   Submitted to EMBL/GenBank/DDBJ, 22-JUL-1992.
XX
DR   EMBL; M11567; AAA51915.1.
DR   PIR; A29565; A29565.
XX
FH   Key             Location/Qualifiers
FT   Region          1..87
FT                   /label=  B chain
FT   Region          90..120
FT                   /label=  C chain
FT   Region          121..187
FT                   /label=  A chain
XX
SQ   Sequence 187 AA;

  Query Match             Score 456;  DB 2;  Length 187;
  Best Local Similarity   Score 456;  Pred. No. 2.5e-43;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86


RESULT 3
AAP40217
ID   AAP40217 standard; protein; 87 AA.
XX
AC   AAP40217;
XX
DT   12-FEB-2002  (first entry)
XX
DE   Sequence of the 42 N-terminal AAs  of preinsulin.
XX
KW   Hormone; cloning vector; phage resistant.
XX
OS   Homo sapiens.
XX
RN   [1]
RP   1-87
RA   Bell G.I.;
RT   ;
RL   Submitted to EMBL/GenBank/DDBJ, 22-JUL-1992.
XX
DR   EMBL; M11567; AAA51915.1.
DR   PIR; A29565; A29565.
XX
FH   Key             Location/Qualifiers
FT   Region          1..87
FT                   /label=  B chain
FT   Region          90..120
FT                   /label=  C chain
FT   Region          121..187
FT                   /label=  A chain
XX
RN   [1]
RP   1-87
RA   (ELIL ) ELI LILLY & CO.;
XX
RT   Recombinant DNA  Escherichia;
XX
DR   WPI; 1984-097703/16.
DR   N-PSDB; AAN10177.
XX
PT   Protecting bacteria from phage infection,  by transformation with
PT   cloning vector contg. segment with restriction and modification
PT   activity.
XX
PS   Example; Fig 1c; 29pp; English.
XX
CC   Plasmid pIN-alpha-1 was digested to insertion a synthetic gene
CC   for thymosin-alpha-1 (AAN10179) into a plasmid pBR322. It is useful in
CC   the construction of plrp24. The inventors claim a method for
CC   protecting bacteria from phage infection  by transformation with
CC   cloning vector contg. segment with restriction and modification
CC   activity. Produn. of plasmid plr 24 is described which uses plrp1 and
CC   produn. of plasmid pBR29 which uses a synthetic gene coding for the
CC   42 N-terminal AAs of preinsulin. (see AAN10179).
XX
SQ   Sequence 87 AA;

  Query Match             Score 456;  DB 2;  Length 87;
  Best Local Similarity   Score 456;  Pred. No. 2.5e-43;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
```

RESULT 7
AAP50000
ID AAP50000 standard; protein; 87 AA.
XX
AC AAP50000;
XX
DT 04-MAR-1998 (first entry)
DT 04-MAR-1998 (last sequence update)
XX
DE Synthetic prohevein.
XX
KW Prohevein-like protein; antifungal; antimicrobial; synthetic.
XX
OS Hevea brasiliensis.
XX
RN [1]
XX
RP 1-87
XX
RA Broekaert W.F., Kelly J., Van Parijs J., Allen A.K., Peumans W.J.;
XX
RT "Antifungal synthetic peptide and variants thereof";
XX
RL Patent number WO9529993-A1/01, 09-NOV-1995;
RL Katholieke Universiteit Leuven (BELG-) Plant Genetic Systems N.V. (PLAN-).
XX
DR WPI; 95-380139/49.
DR N-PSDB; AAN50001.
XX
CC New recombinant DNA expression vector - with promoter,
CC replication and transcription generating polycistronic mRNA
XX
CC Example; Fig 14; 87pp; English.
XX
CC The inventors claim a process for preparing selectable and
CC autonomously replication recombinant DNA expression vectors which
CC comprise 1) a transcriptional and translational activating sequence
CC which is in the reading frame of a nucleotide sequence which codes
CC for a peptide or polypeptide; 2) a translational stop signal; 3) a
CC transcriptional start signal which is in the reading frame of a
CC nucleotide sequence that codes for a fungicidal peptide
CC or 4) an RNA transcript stop signal. The peptide or
CC polypeptide coding sequence codes for less AAs, esp. AAP50001-87.
CC The fungicidal polypeptide is esp. active against fungi, esp.
CC infection and human tissue plasminogen activator.
CC (Updated on 16-AUG-2002 to add missing OS field.)
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 87; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.4e-44;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 1
AAP61990
ID AAP61990 standard; protein; 87 AA.
XX
AC AAP61990;
XX
DT 28-FEB-1994 (first entry)
XX
DE Sequence encoded by the structural gene for human proinsulin.
XX
KW Recombinant plasmid; E.coli expression vector; secretion vector.
XX
OS Homo sapiens.
XX
PN WO9404674-A.
XX
PD 10-MAR-1994.
XX
PF 03-AUG-1992; 92WO-US006287.
XX
PR 05-AUG-1991; 91US-00740287.
PR 28-JAN-1992; 92US-00826607.
XX
PA (GENT) UNIV OF NEW YORK.
XX
PI Inouye M, Nakamura K;
XX
DR WPI; 1994-084182/10.
DR N-PSDB; AAN87425.
XX
PT New recombinant plasmid(s) containing DNA sequences encoding
PT exogenous polypeptide and outer membrane protein of E coli
XX
CC Example; Fig 2; 44pp; English.
XX
CC The invention claims new recombinant plasmids containing DNA sequence
CC which codes a polypeptide, which is foreign to E.coli, in reading phase
CC with a DNA seq. coding for at least one functional fragment derived
CC from an outer membrane lipoprotein gene of E.coli. The foreign gene
CC may be for human insulin. The lipoprotein gene functional fragment
CC may be the promoter, the 5' UTR, the 5' UTR or the transcription
CC termination signal provided that it includes at least the promoter.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 466; DB 7; Length 87;
Best Local Similarity 100.0%; Pred. No. 6.9e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||||||||||||||||||||||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 2
AAP61991
ID AAP61991 standard; protein; 87 AA.
XX
AC AAP61991;
XX
DT 28-FEB-1994 (first entry)
XX
DE Sequence encoded by the structural gene for human proinsulin.
XX
KW Recombinant plasmid; E.coli expression vector; secretion vector.
XX
OS Homo sapiens.
XX
PN WO9404674-A.
XX
PD 10-MAR-1994.
XX
PF 03-AUG-1992; 92WO-US006287.
XX
PR 05-AUG-1991; 91US-00740287.
PR 28-JAN-1992; 92US-00826607.
XX
PA (GENT) UNIV OF NEW YORK.
XX
PI Inouye M, Nakamura K;
XX
DR WPI; 1994-084182/10.
DR N-PSDB; AAN87425.
XX
PT New recombinant plasmid(s) containing DNA sequences encoding
PT exogenous polypeptide and outer membrane protein of E coli
XX
CC Example; Fig 2; 44pp; English.
XX
CC The invention claims new recombinant plasmids containing DNA sequence
CC which codes a polypeptide, which is foreign to E.coli, in reading phase
CC with a DNA seq. coding for at least one functional fragment derived
CC from an outer membrane lipoprotein gene of E.coli. The foreign gene
CC may be for human insulin. The lipoprotein gene functional fragment
CC may be the promoter, the 5' UTR, the 5' UTR or the transcription
CC termination signal provided that it includes at least the promoter.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 466; DB 7; Length 87;
Best Local Similarity 100.0%; Prod. No. 6.9e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||||||||||||||||||||||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

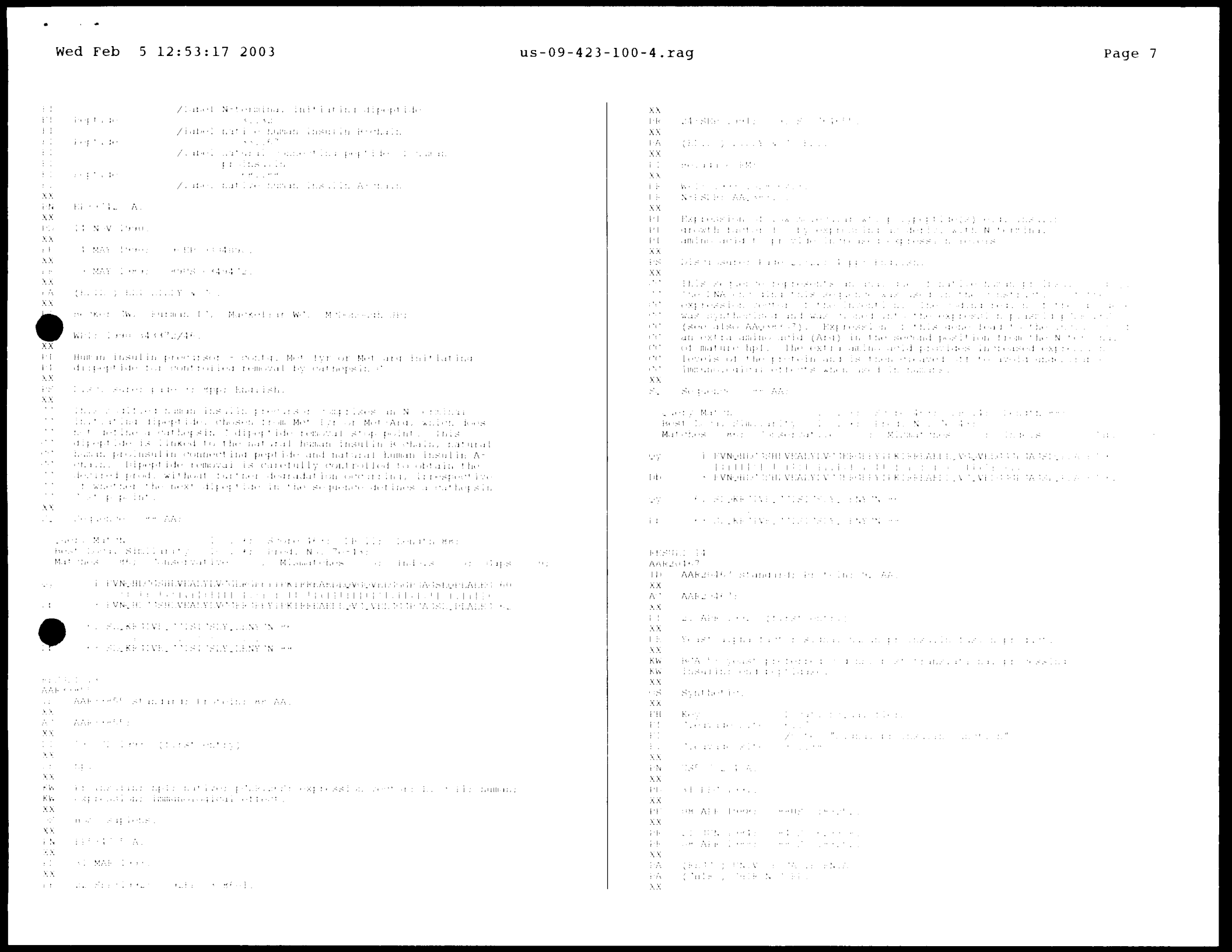
RESULT 3
AAP56767
ID AAP56767 standard; protein; 87 AA.
XX
AC AAP56767;
XX
DT 28-FEB-1994 (first entry)
XX
DE Proinsulin protein sequence.
XX
KW Human; proinsulin; vector; pBCIP; pPINS; rAI; pBE rAI proinsulin;
KW insulin analogue; type I; type II; diabetes.
XX
OS Synthetic.
XX
PN WO9404674-A.
XX
PD 10-MAR-1994.
XX
PF 03-AUG-1992; 92WO-US006287.
XX
PR 05-AUG-1991; 91US-00740287.
PR 28-JAN-1992; 92US-00826607.
XX
PA (PFIZ) PFIZER INC.
PA (SCIO) SCIOS INC.
XX
PI Arad RB, Lucas FR;
XX
DR WPI; 1994-084182/10.
DR N-PSDB; AAN87425.
XX
PT New hepato selective and peripheral selective human insulin
PT analogues and their corresp. DNA, for treatment of type I and
PT type II diabetes.
XX
PS Disclosure; Fig 2; 44pp; English.
XX
CC This sequence represents human proinsulin and was decoded from
CC the sequences given in AAQ36996 7001. The cDNA fragment coding for
CC proinsulin was inserted into plasmid vector pBCIP and digested with
CC KpnI and HindIII. This resulted in the formation of the vector pPINS.
CC A fragment encoding amino acids 1-29 of rAI (see AAQ47002) was inserted
CC into pPINS to give a plasmid which encodes a DNA sequence which is def
CC for amino acids 1-29 of rAI, an 8 amino acid linker sequence and human
CC proinsulin. This plasmid, pBE rAI proinsulin, could be used in the
CC formation of insulin analogues which may be used in the treatment of
CC types I and II diabetes.
XX
SQ Sequence 87 AA;

Query Match 100.0%; Score 466; DB 7; Length 87;
Best Local Similarity 100.0%; Prod. No. 6.9e-43;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
|||||||||||||||||||||||||||
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4
AAP56768
ID AAP56768 standard; protein; 87 AA.
XX
AC AAP56768;
XX
DT 28-FEB-1994 (first entry)
XX
DE Modified human insulin precursor.
XX
KW Human insulin precursor; synthesis; C.
XX
FH Key Location/Qualifiers
FT Peptide 1...



RESULT 1
AAP94[illegible]
ID AAP94[illegible] standard; protein; [illegible] AA.
XX
AC AAP94[illegible];
XX
DT [illegible]-NOV-1989 (first entry)
XX
DE Synthetic proinsulin.
XX
KW Insulin; synthetic gene; yeast.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT source [illegible]
FT /note= "pr insulin chain B"
FT Site 1..[illegible]
FT /note= "dibasic processing site"
FT Peptide [illegible]
FT /note= "proinsulin chain C"
FT Site 1..[illegible]
FT /note= "dibasic processing site"
FT Peptide [illegible]
FT /note= "proinsulin chain A"
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

RESULT 2
AAN[illegible]
ID AAN[illegible] standard; protein; [illegible] AA.
XX
AC AAN[illegible];
XX
PN [illegible]-NOV-1989 (first entry)
XX
PD Expression and secretion of heterologous proteins in yeast using a
PD truncated [illegible] factor leader sequence with a signal peptidase
PD specificity.
XX
PS Preferred: Ala in Z-1 position.
XX
PP The proinsulin is expressed in Saccharomyces cerevisiae AB[illegible].
PP It is encoded by a synthetic gene (see AAN[illegible]). The 3 constituent
PP peptides are proinsulin chains B, C and A resp. The peptides are
PP dibasic processing sites.
XX
SQ Sequence [illegible] AA;

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

Qy [illegible]
Db [illegible]

Search completed: February [illegible], 2003, [illegible]
Job time : [illegible] secs

GenCore version 5.1.8

OM protein - protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 14.1792 Seconds
(without alignments)
583.076 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 86

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	463	100.0	110	1	IPBO	insulin precursor
2	463	100.0	110	2	A42279	insulin precursor
3	456	98.5	110	2	B42279	insulin precursor
4	456	98.5	110	2	A96278	insulin precursor
5	424	91.6	110	1	INPB	insulin precursor
6	417	90.1	110	1	IPBO	insulin precursor
7	404	87.3	86	1	IPBO	insulin precursor
8	404	87.3	110	1	IPBI2	insulin 2 precurso
9	404	87.3	110	1	INMS2	insulin 2 precurso
10	402	86.8	108	2	A96863	insulin precursor
11	402	86.8	110	2	I48266	insulin precursor
12	400	86.4	110	1	IPBII	insulin I precurso
13	384	82.9	84	1	IPBO	insulin precursor
14	369.5	79.8	105	1	IPBO	insulin precursor
15	366	79.0	108	1	INMS1	insulin 1 precurso
16	343.5	74.2	108	2	S06278	insulin precursor
17	329.5	71.2	77	1	INSH	insulin precursor
18	314	67.8	110	1	IPBK	insulin precursor
19	277.5	59.9	109	1	IPBIDO	insulin precursor
20	274.5	59.3	105	2	I51221	insulin precursor
21	265.5	57.4	106	1	IPXL2	insulin II precurs
22	265.5	57.4	107	1	IPCH	insulin precursor
23	262.5	56.7	106	1	IPXL1	insulin I precurso
24	256.5	55.4	51	1	INWHB	insulin - sperm wh
25	256.5	55.4	51	1	INWHF	insulin - finback
26	256.5	55.4	51	1	INEL	insulin - elephant
27	256.5	55.4	81	1	IPDK	insulin precursor
28	256	55.3	96	2	JC0282	epidermal growth f
29	254.5	55.0	51	1	INHY	insulin - hamster
30	251.5	54.3	51	1	INMSSY	insulin - rat (str
31	250.5	54.1	51	2	A42274	insulin precursor
32	246.5	53.2	51	1	INWLS	insulin - sperm wh
33	246.5	53.2	51	1	INSI	insulin - rat
34	246.5	53.2	51	1	INCMA	insulin - guinea-p
35	245.5	53.0	51	1	INCI	insulin - cat
36	244.5	52.8	51	1	INMKSL	insulin - rabbit (
37	239.5	51.7	51	2	JC0282	insulin - Norway r
38	234.5	50.6	51	1	INCB	insulin - bovine (
39	231.5	50.0	51	1	INCS	insulin - sheep
40	227.5	49.1	51	1	INPJ	insulin - porcine
41	227.5	49.1	51	1	INIK	insulin - rabbit (
42	227.5	49.1	51	1	INCS	insulin - rabbit (
43	227.5	49.1	51	2	A61124	insulin - bovine (
44	227.5	49.1	51	2	A61141	insulin - bovine (
45	227	49.0	51	2	S33157	insulin - rabbit (

ALIGNMENTS

RESULT 1
IPBO
insulin precursor [validated] - human
N;Alternate names: preproinsulin
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1981 #sequence_revision 28-Oct-1981 #text_change 16-Jul-1987
C;Accession: A93222; A96255; A93223; A93224; A93225; A93226; A93227; A93228; A
R;Bell, G.I.; Pictet, R.L.; Rutter, W.J.; Cordell, B.; Tischer, E.; Goodman, H.M.
Nature 284, 26-32, 1980
A;Title: Sequence of the human insulin gene.
A;Reference number: A93222; MUID:80127225; PMID:7243470
A;Accession: A93222
A;Molecule type: DNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00265; NID:g186126; PIDN:AAA59172.1; GSPDB:GN00001
R;Ullrich, A.; Bell, J.R.; Chen, E.Y.; Herrera, R.; Petruzzelli, L.M.; Dull, T
Science 209, 612-615, 1980
A;Title: Human insulin gene: nucleotide sequence of preproinsulin and struct
A;Reference number: A96255; MUID:80237225; PMID:6248958
A;Accession: A96255
A;Molecule type: DNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00179; NID:g186125; PIDN:AAA59172.1; GSPDB:GN00001
R;Bell, G.I.; Swain, W.F.; Pictet, R.; Cordell, B.; Goodman, H.M.; Rutter, W.
Nature 282, 525-527, 1979
A;Title: Nucleotide sequence of a DNA coding for human preproinsulin
A;Reference number: A93223; MUID:80047137; PMID:7003284
A;Accession: A93223
A;Molecule type: mRNA
A;Residues: 1-110 <BEL>
A;Cross-references: GB:J00179; NID:g186124; PIDN:AAA59172.1; GSPDB:GN00001
R;Soares, M.; Goodall, G.W.; Bell, G.I.
Science 208, 57-59, 1980
A;Title: Nucleotide sequence of human preproinsulin mRNA
A;Reference number: A96255; MUID:80237225; PMID:6248958
A;Accession: A96255
A;Molecule type: mRNA
A;Residues: 1-110 <SOA>
A;Cross-references: GB:J00265; NID:g186126; PIDN:AAA59172.1; GSPDB:GN00001
R;Nicol, D.S.H.W.; Smith, L.F.
Nature 187, 483-485, 1960
A;Title: Amino-acid sequence of human insulin.
A;Reference number: A93144
A;Accession: A93144
A;Molecule type: protein
A;Residues: 25-54;56-81 <NIC>
R;Boyer, P.E.; Cho, S.; Peterson, J.A.; Steiner, D.F.
J. Biol. Chem. 245, 1577-1580, 1970
A;Title: Studies on human proinsulin. Isolation and partial characterization
A;Reference number: A26075; MUID:71101181; PMID:4911755
A;Accession: A26075
A;Molecule type: protein

A;Residues: 57-87 <OYE>
R;Ko, A.; Smyth, D.G.; Markussen, J.; Sundby, F.
Eur. J. Biochem. 20, 190-199, 1971
A;Title: Amino acid sequence of the C-peptide of human proinsulin.
A;Reference number: A91186; MUID:71257722; PMID:5569404
A;Accession: A91186
A;Molecule type: protein
A;Residues: 57-87 <KOA
R;Bonnefond, A.M.; Julier, C.; Bertrand, J.P.; Bellard, S.; Froguel, P.; Lathrop, M.; Bell
Nature Genet. 4, 305-310, 1993
A;Title: Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment
A;Reference number: 158111; MUID:93364428; PMID:8358443
A;Accession: 158114
A;Status: preliminary, translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-59,63-110 <BES>
A;Cross-references: GB:L15440; NID:g307071; PIDN:AAA59179.1; PID:g307072
R;Sieber, P.; Kamber, B.; Hartmann, A.; Joehl, A.; Riniker, B.; Rittel, W.
Helv. Chim. Acta 57, 2617-2621, 1974
A;Title: Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen.
A;Reference number: A91626; MUID:75077277; PMID:4443293
A;Contents: annotation; synthesis
A;Note: disulfide-bonded human insulin was synthesized; the synthetic hormone was identi
A;Note: article in German with English abstract
R;Nalbandi, V.K.
Hoppe-Seyler's Z. Physiol. Chem. 354, 659-672, 1973
A;Title: The synthesis of C-peptide of human proinsulin.
A;Reference number: A91659; MUID:73040007; PMID:4804504
A;Contents: annotation; synthesis of residues 57-87
R;Geiger, R.; Jaeger, G.; Koenig, W.
Chem. Ber. 106, 2347-2352, 1973
A;Title: Synthesis of the complete sequence of human proinsulin C-peptide and its [Glu
A;Reference number: A90914
A;Contents: annotation; synthesis of residues 57-87
R;Kaufmann, J.E.; Irminger, J.C.; Halban, P.A.
Biochem. J. 310, 869-874, 1995
A;Title: Sequence requirements for proinsulin processing at the B-chain/C-peptide junct
A;Reference number: S58661; MUID:96014185; PMID:7575420
A;Contents: annotation; site-directed mutagenesis study of proteolytic processing
C;Genetics:
A;Gene: GDB:INS
A;Cross-reference: GDB:119849; OMIM:176730
A;Map position: 11p15.5-11p15.5
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54,90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting C-peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,44-109,95-100/Disulfide bonds: #status experimental

  Query Match 100.0%; Score 463; DB 1; Length 110;
  Best Local Similarity 100.0%; Pred. No. 2.9e-43;
  Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 2
A42179
insulin precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A42179; S22058
R;Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-263, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate
A;Reference number: A42179; MUID:92219954; PMID:1560757
A;Accession: A42179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61992; NID:g22202; PIDN:CAA43403.1; PID:g22203
A;Note: sequence extracted from NCBI backbone (NCBIN:95185)

  Query Match 100.0%; Score 463; DB 2; Length 110;
  Best Local Similarity 100.0%; Pred. No. 2.9e-43;
  Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

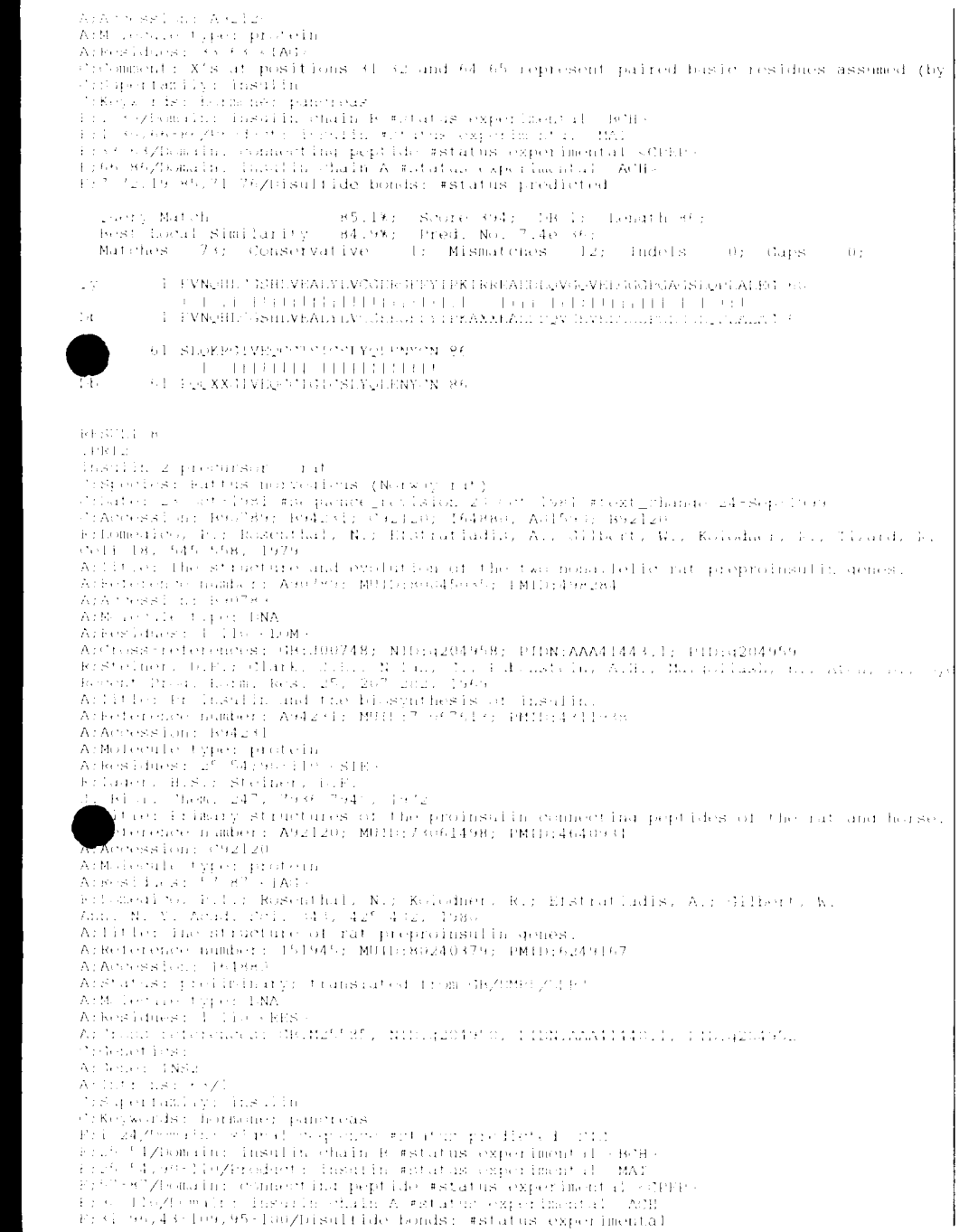
Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 3
B42179
insulin precursor - green monkey
C;Species: Cercopithecus aethiops (green monkey, grivet)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: B42179; A05282; S16494; S22056
R;Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-263, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate
A;Reference number: A42179; MUID:92219954; PMID:1560757
A;Accession: B42179
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61992; NID:g22202; PIDN:CAA43403.1; PID:g22203
A;Note: sequence extracted from NCBI backbone (NCBIN:95185, NCBIP:95194)
R;Peterson, J.D.; Nehrlich, S.; Oyer, P.E.; Steiner, D.F.
J. Biol. Chem. 247, 4866-4871, 1972
A;Title: Determination of the amino acid sequence of the monkey, sheep, and deer proin
A;Reference number: A92111; MUID:72258016; PMID:4626469
A;Accession: A05282
A;Molecule type: protein
A;Residues: 57-87 <PEI>
C;Genetics:
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,44-109,95-100/Disulfide bonds: #status predicted

  Query Match 98.8%; Score 456; DB 2; Length 110;
  Best Local Similarity 98.8%; Pred. No. 1.7e-42;
  Matches 85; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 110

RESULT 4



A;Accession: A04172
A;Molecule type: protein
A;Residues: 1-84 <BRZ>
R;Dodson, G.; Hodgkin, D.; Reynolds, C.
Nature, New Biol. 235, 165-167, 1972
A;Title: The structure of pig and sheep insulins.
A;Reference number: A90344
A;Accession: S16492
A;Molecule type: protein
A;Residues: 1-30;41-51 <BRO>
R;Shoyab, M.; Damsgaard, H.
Hoppe-Seyler's Z. Physiol. Chem. 369, 476-480, 1988
A;Title: Porcine insulin heterogeneity in pigs.
A;Reference number: A68845; MUID:89042178; PMID:3161669
A;Accession: A29855
A;Molecule type: protein
A;Residues: 33-51 <SNE>
A;Note: the authors report the characterization of a connecting peptide variant lacking
A;Accession: A00005
A;Molecule type: protein
A;Residues: 33-51 <SNZ>
R;Goodwin, T.; Dodson, G.; Hodgkin, D.; Mercola, D.
Adv. Protein Chem. 26, 279-362, 1972
A;Title: Insulin, the structure in the crystal and its reflection in chemistry and biolo
A;Reference number: A90017
A;Contents: annotation; X-ray crystallography, 1.9 Angstroms
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status experimental <SIG
F;25-54/Domain: insulin chain B #status experimental <MAT
F;57-84/Domain: connecting peptide #status experimental <CPE
F;89-109/Domain: insulin chain A #status experimental <MAT
F;7,72,89,154/Disulfide bonds: #status experimental

  Query Match 82.7%; Score 383; DB 1; Length 84;
  Best Local Similarity 90.0%; Pred. No. 1.1e-34;
  Matches 71; Conservative 1; Mismatches 9; Indels 2; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKSRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db 61 SLQKRGIVEQCCASVCSLYQLENYCN 84

RESULT 2
INMS1
Insulin I precursor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Apr-1984 #sequence_revision 22-Apr-1995 #text_change 16-Jul-1999
C;Accession: A04172; A00005; A90017; A90344; A68845; A29855; A00006; A00007; A00008; A00
R;Oda-Ameen, Z.; Younes, M.A.; White, J.W.; Beach, E.K.; Floyd, J.C., Jr.; Fineberg, R.N.
Mol. Endocrinol. 1, 327-331, 1987
A;Title: Cloning and nucleotide sequence analysis of complementary deoxyribonucleic acid
A;Reference number: A43090; MUID:88208220; PMID:3453411
A;Accession: A10094
A;Molecule type: mRNA
A;Residues: 1-110 <DAA>
A;Cross-references: GB:M51479; NID:g163478; PIDN:AAA30422.1; PID:g163479
A;Experimental source: fetal pancreas
R;Nolan, C.; Margoliash, E.; Peterson, J.D.; Steiner, D.F.
J. Biol. Chem. 246, 2780-2795, 1971
A;Title: The structure of bovine proinsulin.
A;Reference number: A92080; MUID:71166442; PMID:4928892
A;Accession: A92080
A;Molecule type: protein
A;Residues: 25-109 <NOL>
R;Oyer, P.E.; Horton, R.; Jow, L.; Becker, M.L.; Robertson, A.D.
J. Biol. Chem. 246, 1664-1674, 1971
A;Title: Isolation and characterization of proinsulin C-peptide from bovine pancreas.
A;Reference number: A92071; MUID:71166440; PMID:4928890

A;Accession: A04184
A;Molecule type: protein
A;Residues: 1-82 <SIE>
R;Sakakibara, A.; Smith, L.F.; Matsuzawa, T.; Ohgo, Y.
Eur. J. Biochem. 25, 193-200, 1972
A;Title: Bovine proinsulin: amino acid sequence of the C-peptide and its relation to the
A;Reference number: A91185; MUID:72177211; PMID:5048604
A;Accession: A91185
A;Molecule type: protein
A;Residues: 57-82 <SAK>
R;Sanger, F.; Thompson, E.O.P.
Biochem. J. 53, 366-374, 1953
A;Title: The amino-acid sequence in the glycyl chain of insulin. 2. The investigation
A;Reference number: A90342
A;Accession: A90342
A;Molecule type: protein
A;Residues: 90-105 <SAN>
R;Sanger, F.; Tuppy, H.
Biochem. J. 49, 481-490, 1951
A;Title: The amino-acid sequence in the phenylalanyl chain of insulin. 2. The investig
A;Reference number: A90341
A;Accession: A90341
A;Molecule type: protein
A;Residues: 25-54 <SA2>
R;Chemg, K.; Kawakishi, S.
Eur. J. Biochem. 227, 759-765, 1994
A;Title: Site-specific cleavage of histidine residues in glycine-rich sequences
A;Reference number: S41054; MUID:94178251; PMID:8119944
A;Accession: S41054
A;Molecule type: protein
A;Residues: 25-30 <SHI>
A;Accession: S14164
A;Status: preliminary
A;Molecule type: protein
A;Residues: 25-30,'X',32-42,'X',44-54 <SHZ>
R;Kyte, A.P.; Sanger, F.; Smith, L.F.; Kitai, R.
Biochem. J. 60, 541-556, 1955
A;Title: The disulphide bonds of insulin.
A;Reference number: A90343
A;Contents: annotation; amides; disulfides
R;Wheeler, M.B.; Rutter, W.J.; Sorenson, H.P.; Arvan, P.; Haywood-Reid, P.L.; Yan, Z.; W
FEBS Lett. 316, 295-298, 1993
A;Title: Existence of a proteolytic cofactor that is essential for pancreatic prohormon
A;Reference number: S40258; MUID:93168249; PMID:8431787
A;Accession: S40258
A;Status: preliminary
A;Molecule type: protein
A;Residues: 25-54 <WEN>
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG
F;25-54/Domain: insulin chain B #status experimental <BCH
F;55-56/Region: recognition sequence #status experimental <RAA
F;57-82/Domain: connecting peptide #status experimental <CPE
F;83-108/Domain: insulin chain A #status experimental <ACH
F;31-91,43-108/Disulfide bonds: #status experimental

  Query Match [illegible]; Score [illegible]; DB 2; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
Db 25 FVNQHLCGSHLVEALYLVCGERGFFYTPKAPRDVEDPQVEQLELGGSPGDLQTLALEV 84

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
Db 85 PPQKRGIVEQCCASVCSLYQLENYCN 110

RESULT 3
INMS1
Insulin I precursor - [illegible]

```
 OS  Species: Mus musculus (house mouse)
 DT  Date: 24-Apr-1984 #sequence_revision 14-Jul-1994 #text_change 18-Jun-1999
 AC  Accession: B26842; A48172; A01992; B61012
 RA  Wentworth, B.M.; Schuster, I.M.; Villa-Komaroff, L.; Chirgwin, J.M.
 J. Mol. Evol. 24, 305-312, 1986
 RT  Title: Characterization of the two nonallelic genes encoding mouse preproinsulin.
 RL  Reference number: A92665; MUID:87157763; PMID:3104603
 RA  Accession: B26842
 MT  Molecule type: DNA
 RP  Residues: 1-108 <WEN>
 RX  Cross-references: GB:X04724; NID:g52712; PIDN:CAA28434.1; PID:g52713
 RA  Soares, L.; Cziak, S.; Meriska, H.; Yano, S.
 J. Mol. Endocrinol. 5, 61-67, 1990
 RT  Title: Molecular cloning and DNA sequence analysis of preproinsulin genes in the N=N m
 RL  Reference number: A48172; MUID:90377994; PMID:2407623
 RA  Accession: A48172
 RS  Status: not compared with conceptual translation
 MT  Molecule type: DNA
 RP  Residues: 1-108 <SAW>
 RA  [illegible] H.E.; Gattharr, B.; Kunz, P.; Muellhaupt, E.; Rumbel, R.E.
 RL  Hoppe Seyler's Z. Physiol. Chem. 353, 451-458, 1972
 RT  Title: Amino acid sequence of the two insulins from mouse (Mus musculus).
 RL  Reference number: A01992; MUID:72194455; PMID:5063718
 RA  Accession: A01992
 MT  Molecule type: protein
 RP  Residues: 25-54;88-108 <BOE>
 RA  Bilinski, P.; Nielsen, J.H.; Hansen, B.; Wollinder, B.S.
 J. Chromat. Sci. 4:2, 241-254, 1986
 RT  Title: Reversed-phase high performance liquid chromatographic analyses of insulin bios
 RL  Reference number: A61012; MUID:86202078; PMID:2661585
 RA  Accession: B61012
 MT  Molecule type: protein
 RP  Residues: 57-86 <LIN>
 CC  Function: insulin
 KW  Keywords: hormone; pancreas
 FT  SIGNAL/Domain: signal sequence #status predicted <SIG>
 FT  CHAIN/Domain: insulin chain B #status experimental <MAT>
 FT  PEPTIDE/Domain: connecting peptide #status experimental <CPP>
 FT  CHAIN/Domain: insulin chain A #status experimental <MAT>
 FT  DISULFID: 7-71, 20-96/disulfide bonds: #status predicted

  Query Match          79.0%;  Score 66;  DB 1;  Length 108;
  Best Local Similarity   81.4%;  Pred. No. 1e-32;
  Matches   75;  Conservative    4;  Mismatches   10;  Indels    2;  Gaps    1;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKSRREVEDPQVEQLELGGSPGDLQTLALEV 82

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 86
           |||||||||||||||||||||||||||
Db      83 SLQKRGIVEQCCTSICSLYQLENYCN 108


Search completed: February 5, 2003, 12:45:01
Job time : 5.1792 secs
```

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 PVNLHDLSNHLVEALNLVY...

Scoring table:

Searched: 45 sequences, [illegible] residues

Minimum DB seq length:
Maximum DB seq length:

Post-processing: Minimum Match
Maximum Match
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

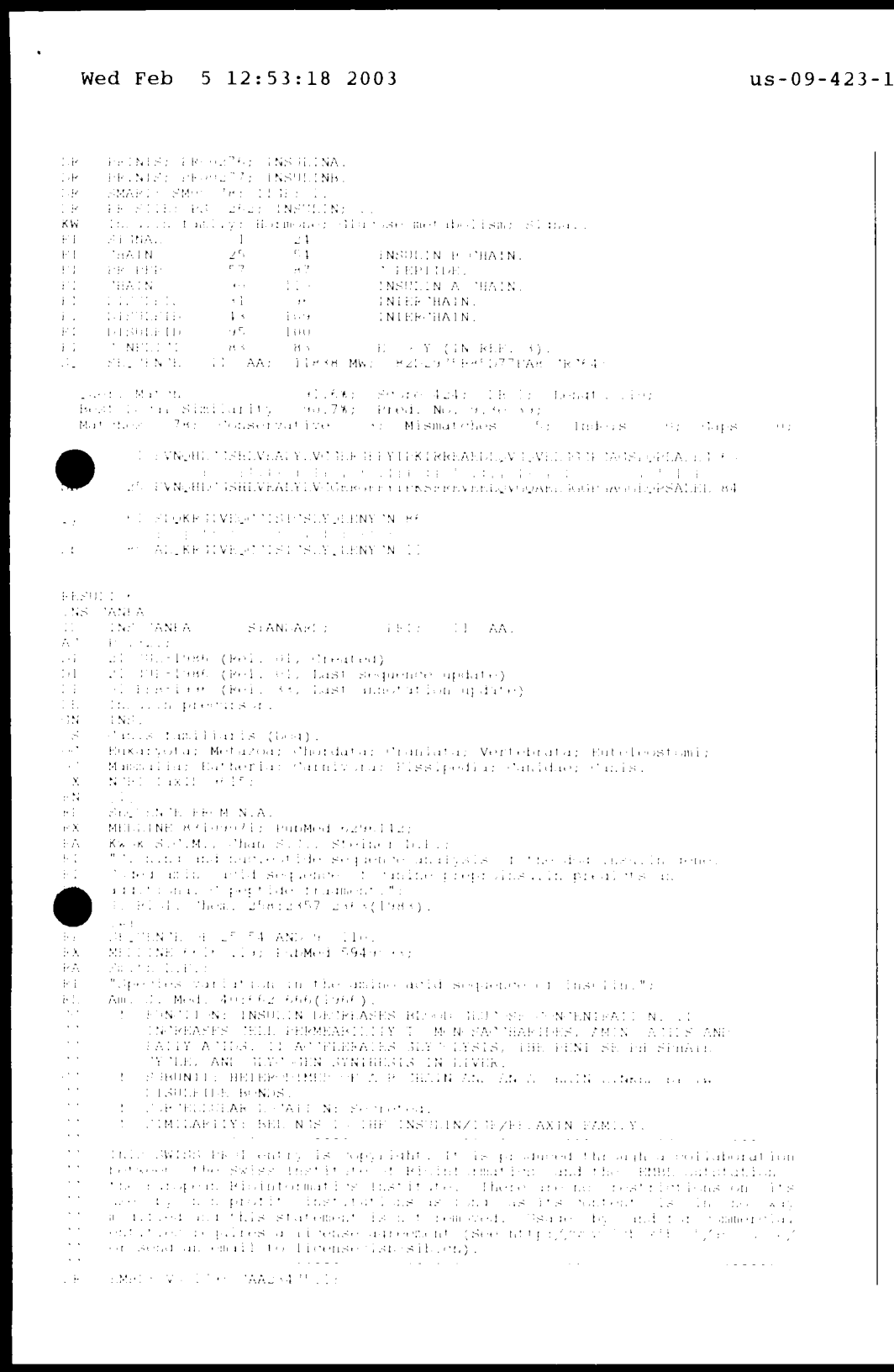
	Score	Match	Length	DB	ID	Description
1	463	100.0	110	1	INS_HUMAN	Insulin homo sapien
2	463	100.0	120	1	INS_PANTR	Insulin pan troglod
3	459	99.1	110	1	INS_CERAE	P30407 cercopithec
4	456	98.5	110	1	INS_MACFA	P30406 macaca fasc
5	424	91.6	110	1	INS_RABIT	P01311 oryctolagus
6	422	91.1	110	1	INS_CANFA	P01321 canis famili
7	394	85.1	86	1	INS_HORSE	P01315 equus caball
8	394	85.1	110	1	INS2_MOUSE	P63326 mus musculu
9	394	85.1	110	1	INS2_RAT	P63323 rattus norv
10	392	84.7	108	1	INS_ALLSI	P14176 alligator m
11	392	84.7	11	1	INS_CHICK	P01323 gallus gall
12	387	83.6	11	1	INS1_RAT	P01322 rattus norv
13	384	82.9	110	1	INS_LAMGL	P01317 lampetra fl
14	384	82.9	110	1	INS_LEPOC	P01316 lepisosteus
15	383	82.7	110	1	INS_PIG	P01320 sus scrofa
16	382	82.5	110	1	INS_BOVIN	P01317 bos taurus
17	380	82.1	108	1	INS1_MOUSE	P63327 mus musculu
18	378.5	81.8	105	1	INS_SHEEP	P01315 ovis aries
19	378.5	81.8	110	1	INS_PIGEO	P22358 columba sp
20	376	81.2	110	1	INS_CHAVI	P14329 chalcalia p
21	375.5	81.1	110	1	INS_[illegible]	P27375 [illegible]
22	375.5	81.1	110	1	INS_SERPE	P27374 scorpaena p
23	375.5	81.1	110	1	INS2_XENLA	P12317 xenopus lae
24	375.5	81.1	110	1	INS_ORNAN	P01322 ornithorhyn
25	374.5	80.9	105	1	INS1_XENLA	P12318 xenopus lae
26	373.5	80.7	110	1	INS_BALPH	P01312 balaenopter
27	373.5	80.7	51	1	INS_ECHMA	P01316 echphis mac
28	373.5	80.7	81	1	INS_ANAPL	P01324 anas platyr
29	373.5	80.7	51	1	INS_ACIBA	P01324 acipenser b
30	373.5	80.7	51	1	INS_BALMU	P01314 balaenopter
31	373.5	80.7	51	1	INS_CAMDR	P01324 camelus dro
32	373.5	80.7	51	1	INS_CAPHI	P01316 capra hircu
33	373.5	80.7	51	1	INS_EQUCA	P01320 equis caball
34	373.5	80.7	51	1	INS_ONCMA	P01324 oncorhynchus m
35	373.5	80.7	51	1	INS_ANSAN	P01321 anser anser
36	372	80.4	110	1	INS_VERM	P01318 [illegible]
37	372.5	80.5	51	1	INS_BOSTA	P01315 [illegible]
38	372.5	80.5	51	1	INS_IBEAT	P01320 [illegible]
39	371	80.2	51	1	INS_ATEG	P01321 [illegible]
40	370.5	80.1	47	1	INS_[illegible]	P01319 [illegible]
41	370.5	80.1	51	1	INS_BOVAN	P01317 [illegible]
42	370.5	80.1	51	1	INS_NOTBE	P01320 [illegible]
43	370.5	80.1	51	1	INS_ALLMI	P01320 [illegible]
44	370	80.0	51	1	INS_[illegible]	P01319 [illegible]
45	370.5	80.1	51	1	INS_ALLMI	P01320 [illegible]

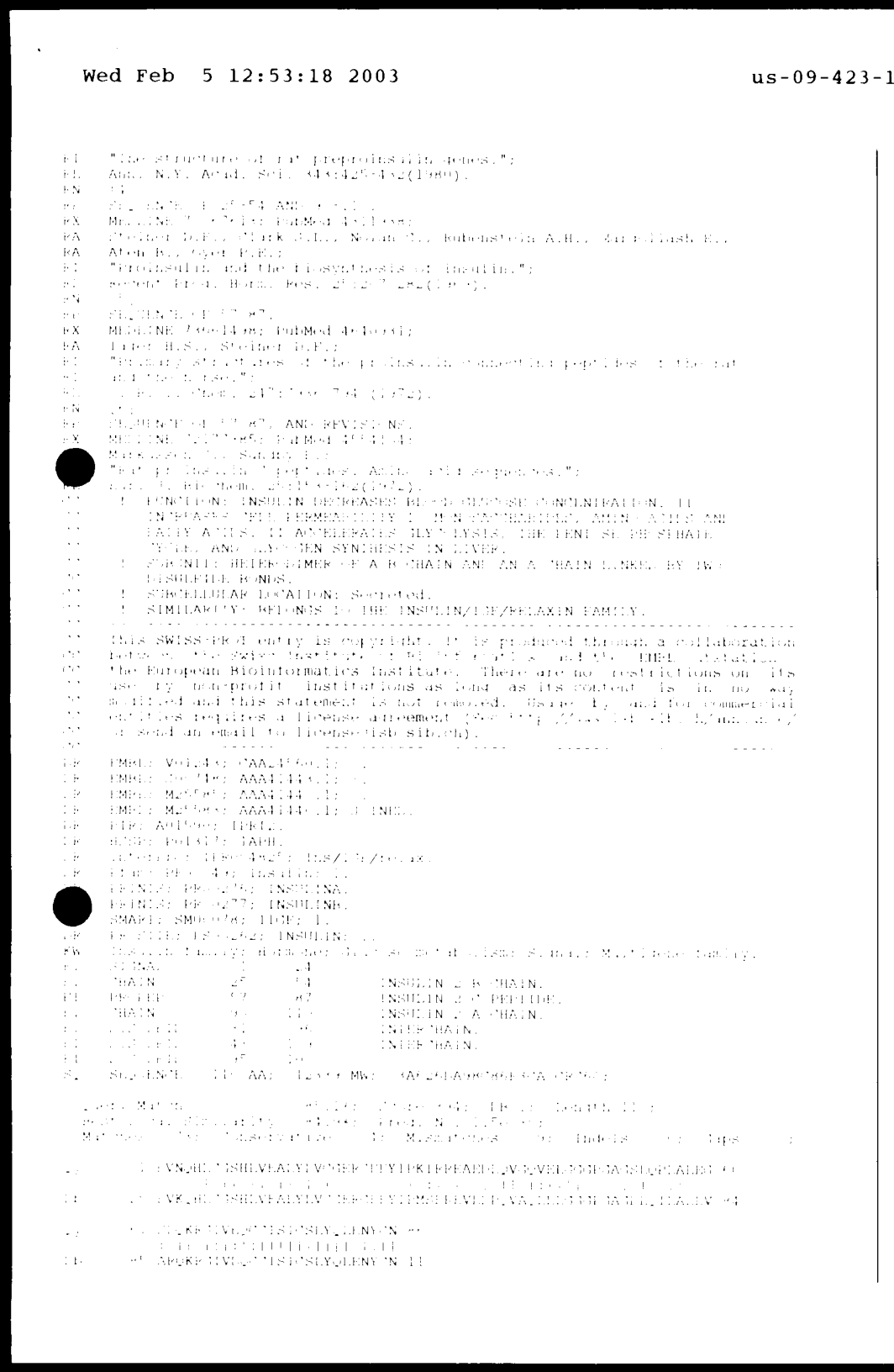
ALIGNMENTS

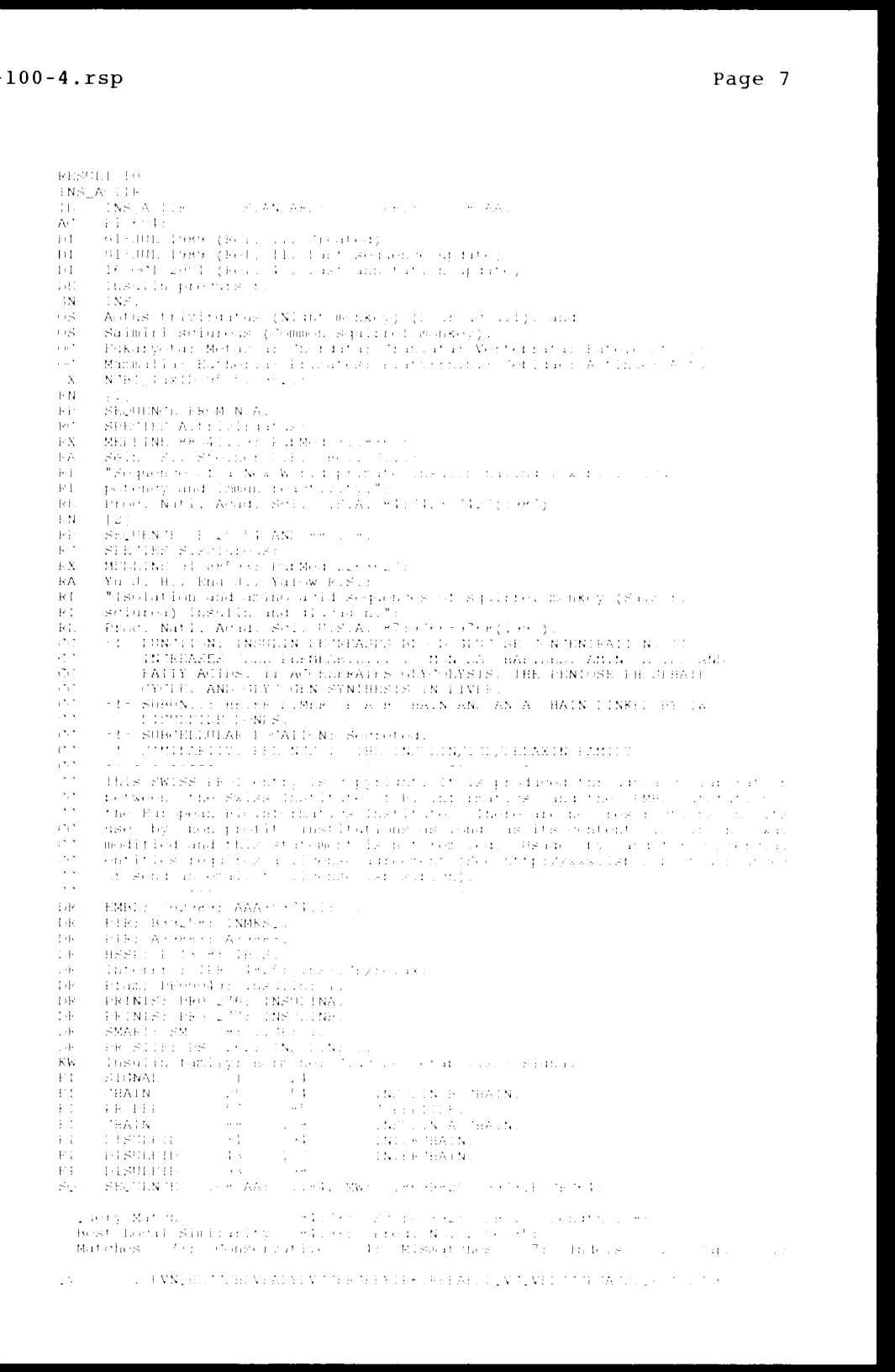
RESULT 1
INS_HUMAN
ID INS_HUMAN STANDARD; PRT; 110 AA.
AC P01308;
DT 21-JUL-1986 (Rel. 1, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 07-JUN-2000 (Rel. 38, Last annotation update)
DE Insulin precursor.
GN INS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80122948; PubMed=6153136;
RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
RA Goodman H.M.;
RT "Sequence of the human insulin gene.";
RL Nature 284:26-32(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157437; PubMed=7012792;
RA Gilbert W., Villa-Komaroff L.;
RT "Useful proteins from recombinant bacteria.";
RL Sci. Am. 242:74-94(1980).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80071165; PubMed=6450083;
RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
RA Rutter W.J.;
RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin";
RL Nature 287:340-342(1980).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157436; PubMed=7012791;
RA Selby M.J., Goeddel D.V., Stebbing N., Yansura D.G.,
RT "Nucleotide sequence of human proinsulin cDNA";
RL Science 209:1244-1246(1980).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157435; PubMed=7012790;
RA Crea R., Kraszewski A., Hirose T., Itakura K.;
RT "Structure of a synthetic human insulin gene";
RL Proc. Natl. Acad. Sci. U.S.A. 77:5765-5769(1980).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157434; PubMed=7012789;
RA Ullrich A., Dull T.J., Gray A., Brownlee G.G.,
RT "Human insulin cDNA: the complete primary structure";
RL DNA 1:1-7(1981).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157433; PubMed=7012788;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157432; PubMed=7012787;
RA Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
RA Goodman H.M.;
RT "Sequence of the human insulin gene.";
RL Nature 284:26-32(1980).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157431; PubMed=7012786;
RA Bell G.I., Swain W.F., Pictet R.L., Cordell B., Goodman H.M.,
RA Rutter W.J.;
RT "Nucleotide sequence of a cDNA clone encoding human preproinsulin";
RL Nature 287:340-342(1980).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157430; PubMed=7012785;
RA Selby M.J., Goeddel D.V., Stebbing N., Yansura D.G.,
RT "Nucleotide sequence of human proinsulin cDNA";
RL Science 209:1244-1246(1980).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157429; PubMed=7012784;
RA Crea R., Kraszewski A., Hirose T., Itakura K.;
RT "Structure of a synthetic human insulin gene";
RL Proc. Natl. Acad. Sci. U.S.A. 77:5765-5769(1980).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157428; PubMed=7012783;
RA Ullrich A., Dull T.J., Gray A., Brownlee G.G.,
RT "Human insulin cDNA: the complete primary structure";
RL DNA 1:1-7(1981).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157427; PubMed=7012782;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [14]
RP TISSUE [illegible];
RA Steiner D.F.;
RL Submitted (MAR-1[illegible]) to the EMBL/GenBank/DDBJ databases.
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157426; PubMed=7012781;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [16]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157425; PubMed=7012780;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [17]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157424; PubMed=7012779;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [18]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157423; PubMed=7012778;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [19]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157422; PubMed=7012777;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [20]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157421; PubMed=7012776;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [21]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157420; PubMed=7012775;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [22]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157419; PubMed=7012774;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [23]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157418; PubMed=7012773;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [24]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157417; PubMed=7012772;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [25]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157416; PubMed=7012771;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [26]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157415; PubMed=7012770;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [27]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157414; PubMed=7012769;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [28]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157413; PubMed=7012768;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).
RN [29]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157412; PubMed=7012767;
RA Efstratiadis A., Posakony J.W., Maniatis T., Lawn R.M.,
RA O'Connell C., Spritz R.A., DeRiel J.K., Forget B.G.,
RA Weissman S.M., Slightom J.L., Blechl A.E., Smithies O.,
RA Baralle F.E., Shoulders C.C., Proudfoot N.J.;
RT "The structure and evolution of the human beta-globin gene family.";
RL Cell 21:653-668(1980).
RN [30]
RP SEQUENCE FROM N.A.
RX MEDLINE=81157411; PubMed=7012766;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA Kolodner R., Tizard R.;
RT "The structure of the human proinsulin gene and the evolution of
RT insulin-like genes";
RL Cell 25:647-654(1981).

RL "...within the B-1 region of insulin gene.";
RL Submitted (04-FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN [8]
RP SEQUENCE OF 25-54 AND 90-110.
RA Nicol T.S.H.W., Scott I.P.;
RT "Amino acid sequence of human insulin.";
RL Nature 244:48-49(1973).
RN [9]
RP SEQUENCE OF 57-87.
RX MEDLINE 71257722; PubMed 5166404;
RA Ko A., Saito T.T., Markussen J., Sundby F.;
RT "The amino acid sequence of the C-peptide of human proinsulin.";
RL Eur. J. Biochem. 20:190-199(1971).
RN [10]
RP SYNTHESIS.
RX MEDLINE 75077277; PubMed 4443299;
RA Sieber P., Kamber B., Hartmann A., Jost A., Sivinker B., Riniker B., Kitzol W.;
RT "Final synthesis of human insulin under directed formation of the
RT disulfide bonds.";
RL Helv. Chim. Acta 57:2617-2621(1974).
RN [11]
RP SYNTHESIS OF 57-87.
RX MEDLINE 75041007; PubMed 4843404;
RA Nilsson V.K.;
RT "Studies on polypeptides. IV. The synthesis of C peptide of human
RT proinsulin.";
RL Hoppe Seyler's Z. Physiol. Chem. 354:669-672(1973).
RN [12]
RP SYNTHESIS OF 65-66 AND 70-74.
RX MEDLINE 74161263; PubMed 4600556;
RA Schlee K., Volk A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 4. Synthesis of the sequences 14-17 and 9-14
RT of human proinsulin C peptides.";
RL Chem. Ber. 106:199-205(1973).
RN [13]
RP SYNTHESIS OF 84-87.
RX MEDLINE 74161261; PubMed 4600553;
RA Schlee K., Lasser L., Krenin W., Trouth L.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 1. Scheme for the synthesis and preparation of
RT the sequence 28-31 of human proinsulin C peptide.";
RL Chem. Ber. 106:188-192(1973).
RN [14]
RP VARIANT LOS ANGELES.
RX MEDLINE 84018058; PubMed 6412455;
RA Hansen H., Chan S.J., Keim P.M., Rubenstein A.H., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB24]insulin.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:6662-6666(1983).
RN [15]
RP VARIANTS LOS ANGELES AND CHICAGO.
RX MEDLINE 84171255; PubMed 6423111;
RA Shoelson S., Fickova M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.;
RT "Identification of a mutant human insulin predicted to contain a
RT basic to neutral alanine substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:7390-7394(1983).
RN [16]
RP VARIANT PROVIDENCE.
RX MEDLINE 87175648; PubMed 3470783;
RA Chan S.J., Seino S., Gruppuso P.A., Schwartz E., Steiner D.F.;
RT "A mutation in the B-chain COOH-terminal region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2194-2198(1987).
RN [17]
RP VARIANT WAKAYAMA.
RX MEDLINE 87[illegible] PubMed [illegible];
RA Sakura H., Iwamoto Y., Sakamoto Y., Kuzuya T., Hirata H.;
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin A8 (Val leu) isolated from the patient.";
RL J. Clin. Invest. [illegible] (1988).
RN [18]
RP VARIANT [illegible].
RX MEDLINE [illegible]; PubMed [illegible];
RA Barbetti F., Raben N., Kadowaki T., Cama A., Accili D., Ostey K.H.,
RA Marynick S.A., Taylor S.I., Roth J.;
RT "Two unrelated patients with identical hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT proinsulin molecule. Identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RL J. Clin. Endocrinol. Metab. 71:1-9(1990).
RN [19]
RP VARIANT ALA-65.
RX MEDLINE 92261096; PubMed 1601997;
RA Shibasaki Y., Kawakami T., Kanazawa Y., Akanuma Y., Takaku F.,
RT "Post-translational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RL J. Clin. Invest. 89:1784-1787(1992).
RN [20]
RP VARIANT KYOTO.
RX MEDLINE 92291497; PubMed 1601997;
RA Yano H., Kitano N., Morimoto M., Nishanky K.S., Imura H., Seino Y.,
RT "A novel point mutation in the human insulin gene (Arg-Leu)
RT hyperproinsulinemia (proinsulin Kyoto).";
RL J. Clin. Invest. 89:1792-1797(1992).
RN [21]
RP STRUCTURE BY NMR.
RX MEDLINE 91194966; PubMed 2271664;
RA Hua Q.-X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 1H NMR
RT assignment of a des pentapeptide analogue and comparison with 2Zn
RT structure.";
RL Biochemistry 29:10541-10555(1990).
RN [22]
RP STRUCTURE BY NMR.
RX MEDLINE 91343567; PubMed 2071[illegible];
RA Hua Q.-X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition.";
RL Biochemistry 30:5505-5515(1991).
RN [23]
RP STRUCTURE BY NMR.
RX MEDLINE 91265527; PubMed 1646835;
RA Hua Q.-X., Weiss M.A.;
RT "Two-dimensional NMR studies of Ins (B23-B27)-beta-Ala: sequence-
RT specific resonance assignments and effects of secondary structure on
RT dynamics.";
RL Biochim. Biophys. Acta 1078:413-426(1991).
RN [24]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Jorgensen A.M.M., Kristensen S.M., Led J.J., Schmidt P.G.;
RT "Three dimensional solution structure of the insulin dimer. A study of
RT the B9(Asp) mutant of human insulin using 2-dimensional proton NMR
RT distance geometry and restrained molecular dynamics.";
RL J. Mol. Biol. 227:1146-1163(1992).
RN [25]
RP STRUCTURE BY NMR OF VARIANT LOS ANGELES.
RX MEDLINE [illegible]; PubMed 8[illegible];
RA Hua Q.-X., Shoelson S.E., Inouye K., Weiss M.A.;
RT "Paradoxical structure and function in a mutant human insulin
RT associated with diabetes mellitus.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] (1995).
RN [26]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed 9[illegible];
RA Chang X., Jorgensen A.M., Bardram P., Led J.J.;

```
RL   (or EC)       87     87       Eliminate.
FT   CHAIN         90    110       INSULIN A CHAIN.
FT   DISULFID      91     96       INTERCHAIN.
FT   DISULFID      95    107       INTERCHAIN.
FT   CONFLICT      95     95
SQ   SEQUENCE   110 AA;  12015 MW;  95A11F4B8E2L427P CRC64;

  Query Match             98.5%;  Score 756;  DB 1;  Length 110;
  Best Local Similarity   98.8%;  Pred. No. 3.4e-42;
  Matches   85;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60

Qy      61 LQKRGIVEQCCTSICSLYQLENYCN 86
           |||||||||||||||||||||||||
Db      61 LQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 3
INS_MACFA
ID   INS_MACFA      STANDARD;      PRT;   110 AA.
AC   P01320;
DT   21-JUL-1986 (Rel. 01, Created)
DT   15-AUG-1987 (Rel. 05, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88208174; PubMed=3124262;
RA   Wetekam W., Gruendeken J., Schneweber M., Wandelmayer F.,
RA   Winter-Soer E.;
RT   "Structure of the sequence of DNA coding for preproinsulin from the
RT   primate Macaca fascicularis.";
RL   Gene 61:79-84(1987).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X14966; CAA33004.1;
DR   PIR; A29778; A29778.
DR   HSSP; P01308; 1AIO.
DR   Interpro; IPR000538; Ins/IGF/relax.
DR   Pfam; PF00049; insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal.
FT   SIGNAL        1     24
FT   CHAIN        25     54       INSULIN B CHAIN.
FT   PROPEP       57     87       C-PEPTIDE.
FT   CHAIN        90    110       INSULIN A CHAIN.
FT   DISULFID     31     96       INTERCHAIN.
FT   DISULFID     95    107       INTERCHAIN.
SQ   SEQUENCE   110 AA;  12015 MW;  [illegible] CRC64;

  Query Match             98.5%;  Score 756;  DB 1;  Length 110;
  Best Local Similarity   98.8%;  Pred. No. 3.4e-42;
  Matches   85;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60

Qy      61 LQKRGIVEQCCTSICSLYQLENYCN 86
           |||||||||||||||||||||||||
Db      61 LQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4
INS_RABIT
ID   INS_RABIT      STANDARD;      PRT;   110 AA.
AC   P01317;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Pancreas;
RX   MEDLINE=81170459; PubMed=7012449;
RA   Bell G.I., Pictet R.L., Rutter W.J., Cordell B., Tischer E.,
RA   Goodman H.M.;
RT   "Sequence of the human insulin gene.";
RL   Nature 284:26-32(1980).
RN   [2]
RP   SEQUENCE OF 56-110 FROM N.A.
RA   Tobkes N.J., Wheeler C.E., Permutt M.A.;
RL   Submitted (JUN-1983) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
CC
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X14966; AAA31954.1;
DR   EMBL; M21154; AAA31540.1;
DR   PIR; A23611; INRB.
DR   HSSP; P01308; 1IVM.
DR   Interpro; IPR000538; Ins/IGF/relax.
DR   Pfam; PF00049; insulin; 1.
```

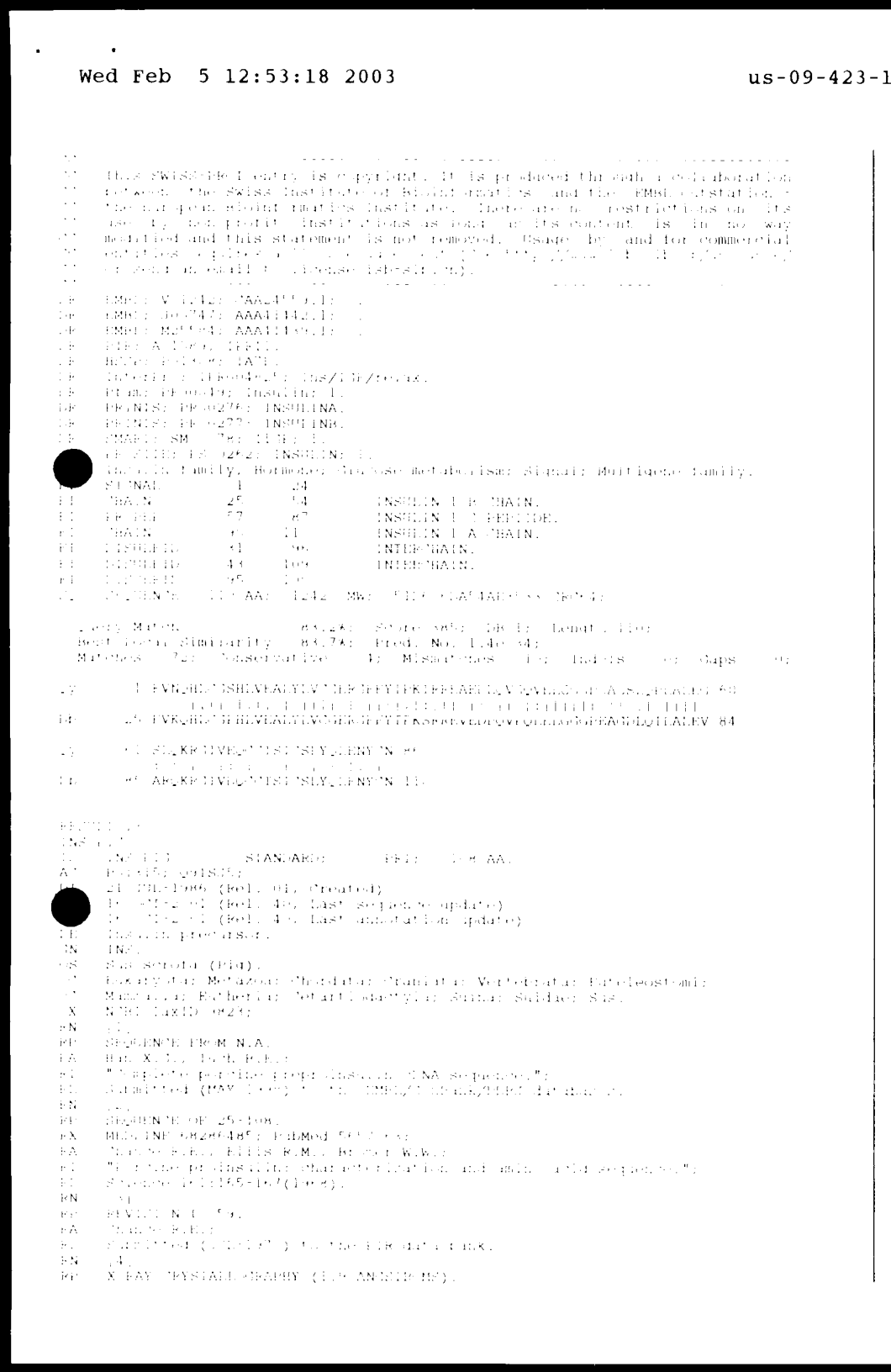






```
ID   INS_MESAU               STANDARD;      PRT;   110 AA.
AC   P01315;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-JAN-1990 (Rel. 18, Last sequence update)
DT   01-FEB-1995 (Rel. 28, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Mesocricetus auratus (Golden hamster) (Syrian hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Cricetulus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84110646; PubMed=6690663;
RA   Bell G.I., Sanchez-Pescador R.;
RT   "Sequence of a DNA coding for Syrian hamster preproinsulin.";
RL   Diabetes 33:207-209(1984).
RN   [2]
RP   SEQUENCE OF 25-54 AND 90-110.
RA   Neon C.F.A., Belcher H.K., Steiner H., Lebovitz H.E.;
RT   "Structural Hamster Insulin: Comparison with a Human Insulin.";
RL   Fed. Proc. 33:600-606(1974).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; M22408; AAA37089.1; -.
DR   PIR; A24145; INRY.
DR   HSSP; 1ins; 1IYM.
DR   InterPro; IPR003235; Insulin_rel.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal.
FT   SIGNAL       1     24
FT   CHAIN       25     54       INSULIN B CHAIN.
FT   PROPEP       55     89       C PEPTIDE.
FT   CHAIN        90    110       INSULIN A CHAIN.
FT   DISULFID     31     96       INTERCHAIN.
FT   DISULFID     43    109       INTERCHAIN.
FT   DISULFID     96    101
SQ   SEQUENCE   110 AA;  12260 MW;  21D892B8CA5C0B3F CRC64;

  Query Match             81.7%;  Score 494;  DB 1;  Length 110;
  Best Local Similarity   84.9%;  Pred. No. 2.5e-46;
  Matches   73;  Conservative    4;  Mismatches    9;  Indels    4;  Gaps    1;
```

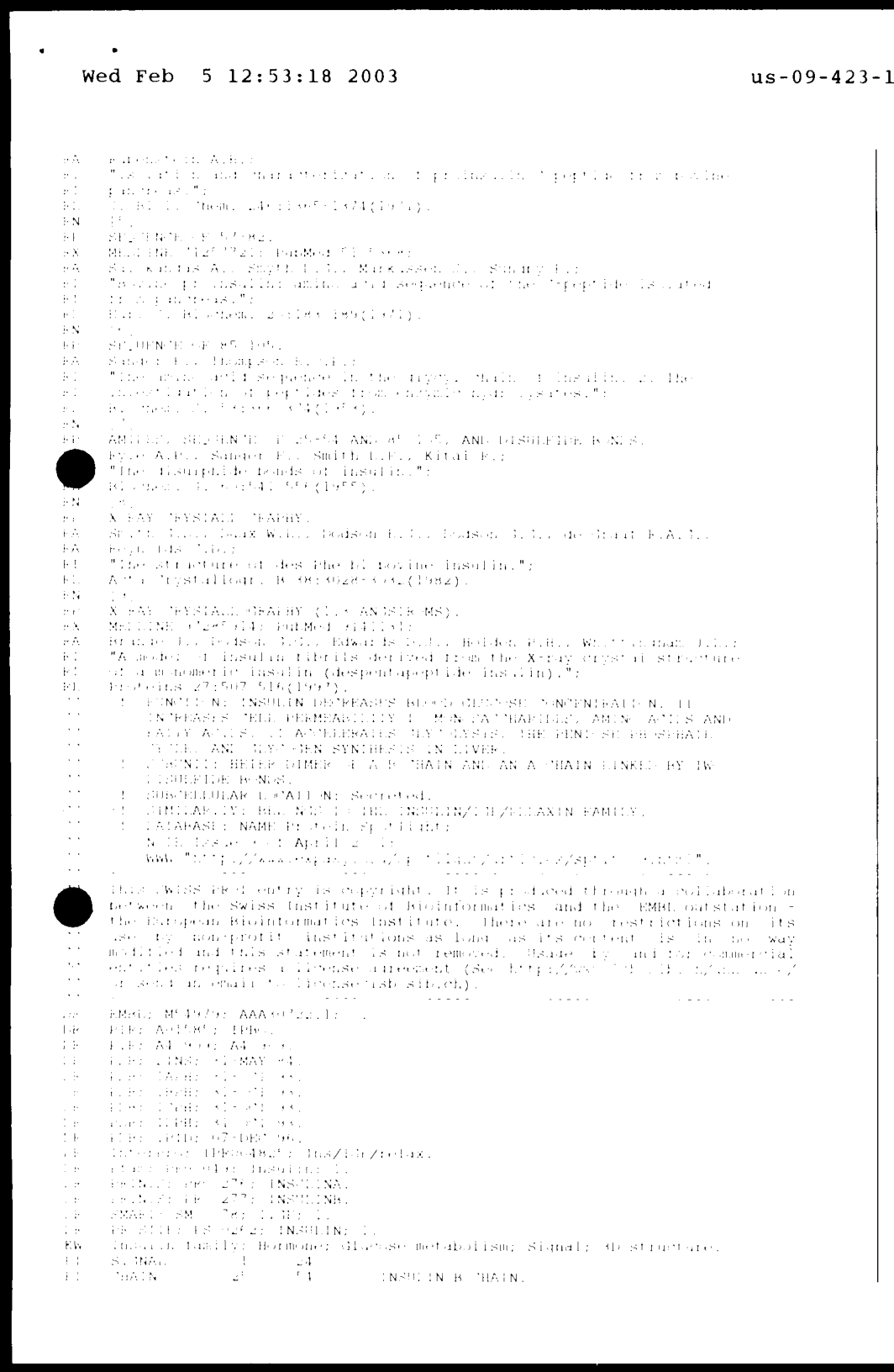
```
ID   INS_RAT                 STANDARD;      PRT;   110 AA.
AC   P01325;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Insulin I precursor.
GN   INSI OR INS 1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=80041044; PubMed=6997466;
RA   Soares M.B., Bon R., Jeffrey J., Chirgwin J.M., Bell G.I.;
RA   Efstratiadis A., Ruttner W., Maniatis T.;
RT   "Isolation and characterization of a cloned rat insulin gene.";
RL   Cell 18:545-558(1979).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=80215027; PubMed=6248284;
RA   Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,
RA   Tizard R.;
RT   "The structure and evolution of the two nonallelic rat preproinsulin
RT   genes.";
RL   Cell 18:545-558(1979).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Brown-VANE B02403?9; PubMed 6249167;
RA   Lomedico P.T., Rosenthal N., Kolodner R., Efstratiadis A.,
RA   Gilbert W.;
RT   "The structure of rat preproinsulin genes.";
RL   Ann. N.Y. Acad. Sci. 358:425-432(1980).
RN   [4]
RP   SEQUENCE OF 25-54 AND 90-110.
RX   MEDLINE ?-?????; PubMed 6435???;
RA   Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margoliash E.,
RA   Aten B., Oyer P.E.;
RT   "Proinsulin and the biosynthesis of insulin.";
RL   Recent Prog. Horm. Res. 25:??? ???-???(1969).
RN   [5]
RP   SEQUENCE OF 57-87.
RX   MEDLINE ?3?14??; PubMed 4?4????;
RA   Jamet H.P., Steiner D.F.;
RT   "Primary structures of the proinsulin connecting peptides from rat
RT   and the horse.";
RL   J. Biol. Chem. 247:7930-7941(1972).
RN   [6]
RP   SEQUENCE OF 57-87, AND REVISIONS.
RX   MEDLINE 72127???; PubMed 4554164;
RA   Markussen J., Sundby F.;
RT   "Rat preproinsulin 2 peptides. Amino acid sequences.";
RL   Eur. J. Biochem. 27:153-162(1972).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
```



```
ID   1IRS_BOVIN  STANDARD;  PRT;  110 AA.
AC   P01317;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1992 (Rel. 25, Last sequence update)
DT   15-OCT-2002 (Rel. 41, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86280209; PubMed=3516742;
RA   Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA   Kolodner R., Tizard R.;
RT   "The structure of the bovine insulin gene and its mRNA.";
RL   Cell 40:377-384(1985).
RN   [2]
RP   SEQUENCE OF 24-110.
RX   MEDLINE=71164442; PubMed=4900061;
RA   Neurath H., Mansfield S.G., Peterson J.A., Steiner D.F.;
RT   "The structure of bovine proinsulin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:2870-2873(1971).
RN   [3]
RP   SEQUENCE OF 24-54.
RA   Sanger F., Tuppy H.;
RT   "The amino-acid sequence in the phenylalanyl chain of insulin. ...the
RT   investigation of peptides from enzymic hydrolysates.";
RL   Biochem. J. 49:481-490(1951).
RN   [4]
RP   SEQUENCE OF 57-82.
RX   MEDLINE=71164443; PubMed=4900062;
RA   Steiner D.F., Cho S.Y., Pilkis S.J., Peterson J.A.,
```

```
CC   modified and this Statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -!- ----------- ---- - --
DR   EMBL; X56411; CAA39971.1; -.
DR   PIR; I49078; IAIS.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal; 3D-structure.
FT   SIGNAL        1     24       BY SIMILARITY.
FT   CHAIN        25     54       INSULIN B CHAIN.
FT   PROPEP       57     87       C PEPTIDE.
FT   CHAIN        90    110       INSULIN A CHAIN.
FT   DISULFID     31     96       INTERCHAIN.
FT   DISULFID     43    107       INTERCHAIN.
FT   DISULFID     91    103
FT   HELIX        29     34
FT   STRAND       38     38
FT   HELIX        40     53
FT   HELIX       100    107
FT   STRAND      109    110
SQ   SEQUENCE   110 AA;  11671 MW;  B44C1B12C057EBBC CRC64;

  Query Match             81.4%;  Score 747;  DB 1;  Length 110;
  Best Local Similarity   81.4%;  Pred. No. 2.2e-54;
  Matches   74;  Conservative    1;  Mismatches    5;  Indels    2;  Gaps    1;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR 64
           |||||||||||||||||||||||||||||||||||| || || |||||| |||| || |||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKR 84

Qy      65 GIVEQCCTSICSLYQLENYCN 85
           |||||||||| |||| |||||
Db      85 GIVEQCCTSICSLYQLENYCN 105
```

```
RESULT 15
INS_BOVIN
ID   INS_BOVIN      STANDARD;      PRT;   110 AA.
AC   P01317;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1992 (Rel. 25, Last sequence update)
DT   15-OCT-2002 (Rel. 41, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86280209; PubMed=3516742;
RA   Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W.,
RA   Kolodner R., Tizard R.;
RT   "The structure of the bovine insulin gene and its mRNA.";
RL   Cell 40:377-384(1985).
RN   [2]
RP   SEQUENCE OF 24-110.
RX   MEDLINE=71164442; PubMed=4900061;
RA   Neurath H., Mansfield S.G., Peterson J.A., Steiner D.F.;
RT   "The structure of bovine proinsulin.";
RL   Proc. Natl. Acad. Sci. U.S.A. 68:2870-2873(1971).
RN   [3]
RP   SEQUENCE OF 24-54.
RA   Sanger F., Tuppy H.;
RT   "The amino-acid sequence in the phenylalanyl chain of insulin. ...the
RT   investigation of peptides from enzymic hydrolysates.";
RL   Biochem. J. 49:481-490(1951).
RN   [4]
RP   SEQUENCE OF 57-82.
RX   MEDLINE=71164443; PubMed=4900062;
RA   Steiner D.F., Cho S.Y., Pilkis S.J., Peterson J.A.,
```



Run on: February 5, 2003, 12:42:21 ; Search time 0.542 Seconds
(without alignments)
631.727 Million cell updates/sec

Title: **US-09-423-100-4**
Perfect score: 463
Sequence: 1 FVNQHLCGSHLVEALYLVC.....IVEQCCTSICSLYQLENYCN 49

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1115 seqs, 2347117 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : **SPTREMBL_21:***
1: **sp_archea:***
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus2:*
16: sp_bacteria2:*
17: sp_archea2:*

Total No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB	ID	Description
1	114	80.2	11	1	Q6XIX1 Spermophilus
2	108	75.9	11	1	Q9WNW7 Ictidomys tri
3	107.5	74.4	11	1	Q9TA47 Otocerlobus
4	107.5	74.4	11	1	Q9TAA8 Potamotrygon
5	106.5	74.1	11	1	Q9DAY1 Rhinolophus
6	105	47.4	11	1	Q9MTB0 Chitala chi
7	104.5	73.3	10	1	Q9TIE5 Eschrichtiu
8	103.5	71.9	1	1	Q9JKN1 Mus musculu
9	103.5	71.9	11	1	Q9TA47 Balaenoptera
10	103.5	44.4	1	1	Q9TBC1 Ratufa bada
11	103.5	44.4	11	1	Q9TBB2 Ornithorhync
12	103	46.0	11	1	Q9Z543 Mus spretus
13	102	45.1	11	1	Q9Z542 Mus spretus
14	101.5	71.9	15	1	Q9TBG7 Paradelphis
15	101.5	71.9	10	1	Q7YBZ4 Orycteropus

ALIGNMENTS

RESULT 1
Q6XIX1
ID Q6XIX1 PRELIMINARY; PRT; 49 AA.
AC Q6XIX1;
DT 01-DEC-2004 (TrEMBLrel. 24, Created)
DT 01-DEC-2004 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2005 (TrEMBLrel. 27, Last annotation update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciuromorpha; Sciuridae; Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Hoshino M.M., Kock M.E., Schandl C.A., Spirito F.E., Adams W.J.;
RT "Regulation of ECRG expression in a hibernating mammal.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: NOT SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AY496644; AAK61544.1; -.
DR InterPro; IPR004825; Ins/IGF/Rel.fac.
DR Pfam; PF00049; Ins_IGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 49 AA; 5 MW; 0000000000000000 CRC64;

Query Match 100.0%; Score 114; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 49

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 49

RESULT 2
Q6XIX1
ID Q6XIX1 PRELIMINARY; PRT; 49 AA.
AC Q6XIX1;
DT 01-DEC-2004 (TrEMBLrel. 24, Created)
DT 01-DEC-2004 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2005 (TrEMBLrel. 27, Last annotation update)
DE Insulin.
OS Spermophilus tridecemlineatus (Thirteen-lined ground squirrel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciuromorpha; Sciuridae; Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Hoshino M.M., Kock M.E., Schandl C.A., Spirito F.E., Adams W.J.;
RT "Regulation of ECRG expression in a hibernating mammal.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: NOT SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AY496644; AAK61544.1; -.
DR InterPro; IPR004825; Ins/IGF/Rel.fac.
DR Pfam; PF00049; Ins_IGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 49 AA; 5 MW; 0000000000000000 CRC64;

Query Match 100.0%; Score 114; DB 11; Length 49;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 49

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 49

RESULT 1
Q9WNW6
ID Q9WNW6 PRELIMINARY; PRT; 71 AA.
AC Q9WNW6;
DT 01-MAR-2000 (TrEMBLrel. 16, Created)
DT 01-MAR-2000 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin.
GN Felis silvestris catus (Cat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Kato Y., Morimatsu M.;
RT "Cat insulin.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB043545; BAB84110.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; UNKNOWN_1.
SQ SEQUENCE 71 AA; 7990 MW; 95FB0217A77BED7A1 CRC64;

Query Match 96.2%; Score 708; DB 5; Length 71;
Best Local Similarity 98.7%; Pred. No. 1.9e-39;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 71

RESULT 2
Q9TIB7
ID Q9TIB7 PRELIMINARY; PRT; 106 AA.
AC Q9TIB7;
DT 01-MAR-2000 (TrEMBLrel. 16, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin.
GN Panthera pardus (Northern leopard frog).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20625507; PubMed=11019774;
RA Irwin D.M., Sivarajah P.;
RT "Preproinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
RT proteolytic processing.";
RL Comp. Biochem. Physiol. 125B:465-471(2000).
DR EMBL; AF252187; AAF87208.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 106 AA; 12187 MW; 8A97DEB97C217FD2 CRC64;

Query Match 96.2%; Score 708; DB 1; Length 106;
Best Local Similarity 98.7%; Pred. No. 2.8e-39;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 35 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 106

RESULT 3
Q9TIB6
ID Q9TIB6 PRELIMINARY; PRT; 106 AA.
AC Q9TIB6;
DT 01-MAR-2000 (TrEMBLrel. 16, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Preproinsulin.
GN Pantherophis guttatus (Corn snake) (Elaphe guttata).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20625507; PubMed=11019774;
RA Irwin D.M., Sivarajah P.;
RT "Preproinsulin cDNAs from the leopard frog, Rana pipiens: evolution of
RT proteolytic processing.";
RL Comp. Biochem. Physiol. 125B:465-471(2000).
DR EMBL; AF252187; AAF87208.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 106 AA; 12187 MW; 8A97DEB97C217FD2 CRC64;

Query Match 96.2%; Score 708; DB 1; Length 106;
Best Local Similarity 98.7%; Pred. No. 2.8e-39;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 35 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 106

RESULT 4
Q98IA7
ID Q98IA7 PRELIMINARY; PRT; 71 AA.
AC Q98IA7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproinsulin (Fragment).
GN Cynoglossus semilaevis (Tongue sole).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Heteromorpha;
OC Pleuronectiformes; Cynoglossidae; Cynoglossinae.
OX NCBI_TaxID=107571;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21178477; PubMed=11430717;
RA Xu H., Duan Z., Li Y., Xu Q., Wang X., Graham A., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteichthyan morphs and a cyprinid.";
RL Mol. Cell. Endocrinol. 174:51-58(2001).
DR EMBL; AF338422; AAK38777.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Xu H., Duan Z., Li Y., Xu Q., Wang X., Graham A., Youson J.H.;
RT "Molecular cloning of preproinsulin cDNAs from several
RT osteichthyan morphs and a cyprinid.";
RL Mol. Cell. Endocrinol. 174:51-58(2001).
DR EMBL; AF338423; AAK38778.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 71 AA; 7990 MW; A79EE0D2D4B0DB20 CRC64;

Query Match 96.2%; Score 708; DB 2; Length 71;
Best Local Similarity 98.7%; Pred. No. 1.9e-39;
Matches 72; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYMVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 71

RESULT 5
Q98IA8
ID Q98IA8 PRELIMINARY; PRT; 110 AA.
AC Q98IA8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproinsulin.
GN Pantodon buchholzi (Butterfly fish).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteoglossomorpha;
OC Osteoglossidae; Pantodontidae; Pantodon.
OX NCBI_TaxID=8276;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21178477; PubMed=11430717;

```
RESULT 12
Q08182
ID   Q08182      PRELIMINARY;      PRT:    91 AA.
AC   Q08182;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Preproinsulin (Fragment).
OS   Amblyoplites ingestris.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Perciformes; Percoidei;
OC   Centrarchidae; Amblopliti.
OX   NCBI_TaxID=13273;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Al-Mahrouki A.A., Irwin L.M., Youson J.H.;
RL   "Molecular cloning of preproinsulin cDNA from the rock bass.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
DR   EMBL; AF149581; AAK28708.1; -.
DR   HSSP; P01308; 1LPH.
DR   InterPro; IPR004061; Insulin_chordata.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         2      91       A CHAIN.
SQ   SEQUENCE   91 AA;  9961 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 203.5;  DB 13;  Length 91;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-15;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       8 LHLVGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALE 59
           |||||||:|||||||||:|||| || ||:|:|      |:|    ||    :   |:|
Db       4 LHLCGSHLVDALYLVCGDRGFFYNPKDDIHRRSVEQPLALESGLQKRGIVEQCCTS 61

Qy      60 SLQKRGIVEQCCTSICSLYQLENYCN 86
           ||   |||||||||  |:::|  |||
Db      62 LQMEMMVKRGIVEQCCHHPCNLFELKNYCN 91


RESULT 13
Q62542
ID   Q62542      PRELIMINARY;      PRT:    47 AA.
AC   Q62542;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Insulin 2 (Fragment).
GN   INS2.
OS   Mus spretus (Western wild mouse)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STRAIN:CHEEU/EI;
RX   MEDLINE=94319092; PubMed=8041640;
RA   Ko M.S., Wen X., Horton J.D., Kelly M.L., Takahashi N., Maezaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL   Mamm. Genome 5:349-355(1994).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; U05729; AAB06434.1; -.
DR   HSSP; P01308; 1ZNI.
DR   MGD; MGI:96572; Ins2.
DR   InterPro; IPR004061; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         2      47       A CHAIN.
SQ   SEQUENCE   47 AA;  5161 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 203.5;  DB 13;  Length 47;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-15;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       8 LHLVGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALE 59
           |||||||:|||||||||:|||| || ||:|:|      |:|    ||    :   |:|
Db       4 LHLCGSHLVDALYLVCGDRGFFYNPKDDIHRRSVEQPLALESGLQKRGIVEQCCTS 61

Qy      60 SLQKRGIVEQCCTSICSLYQLENYCN 86
           ||   |||||||||  |:::|  |||
Db      62 LQMEMMVKRGIVEQCCHHPCNLFELKNYCN 91


RESULT 14
Q62542
ID   Q62542      PRELIMINARY;      PRT:    47 AA.
AC   Q62542;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Insulin 2 (Fragment).
GN   INS2.
OS   Mus spretus (Western wild mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10096;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   STRAIN:CHEEU/EI;
RX   MEDLINE=94319092; PubMed=8041640;
RA   Ko M.S., Wen X., Horton J.D., Kelly M.L., Takahashi N., Maezaki Y.,
RA   Nadeau J.H.;
RT   "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RL   Mamm. Genome 5:349-355(1994).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   EMBL; U05729; AAB06434.1; -.
DR   HSSP; P01308; 1ZNI.
DR   MGD; MGI:96572; Ins2.
DR   InterPro; IPR004061; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Hormone; Glucose metabolism; Multigene family.
FT   NON_TER       1       1
FT   CHAIN         2      47       A CHAIN.
SQ   SEQUENCE   47 AA;  5161 MW;  [illegible] CRC64;

  Query Match             44.0%;  Score 203.5;  DB 13;  Length 47;
  Best Local Similarity   46.7%;  Pred. No. 1.5e-15;
  Matches   42;  Conservative   13;  Mismatches   26;  Indels    9;  Gaps    4;

Qy       8 LHLVGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALE 59
           |||||||:|||||||||:|||| || ||:|:|      |:|    ||    :   |:|
Db       4 LHLCGSHLVDALYLVCGDRGFFYNPKDDIHRRSVEQPLALESGLQKRGIVEQCCTS 61

Qy      60 SLQKRGIVEQCCTSICSLYQLENYCN 86
           ||   |||||||||  |:::|  |||
Db      62 LQMEMMVKRGIVEQCCHHPCNLFELKNYCN 91
```

```
DB       1 GSP--GDQGILALDVAGKRSCIVLGGTGTGKSLVGLGNYPN 39


RESULT 15
P93607
ID   P93607      PRELIMINARY;      PRT;   159 AA.
AC   P93607;
DT   01-NOV-1998 (TrEMBLrel. 03, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Preproinsulin-like growth factor Ia.
GN   IGF-1.
OS   Paralichthys olivaceus (Flounder).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC   Pleuronectoidei; Paralichthyidae; Paralichthys.
OX   NCBI_TaxID=8255;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim I.S.;
RT   "Expression of IGF Ib cDNA clone isolated from Paralichthys olivaceus
RT   in mammalian CHO cell line using green fluorescence protein (GFP)
RT   tagging: secretory production of big IGFIb-GFP fusion proteins from
RT   stable transfected CHO cell culture.";
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
DR   EMBL; AJ010602; CAA09267.1; -.
DR   HSSP; P01343; 1GFL.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULIND.
DR   SMART; SM00078; 1IGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
SQ   SEQUENCE   159 AA;  17511 MW;  3B9C1C8B29B3EB3B CRC64;

  Query Match             30.8%;  Score 142.5;  DB 13;  Length 159;
  Best Local Similarity   37.5%;  Pred. No. 4.3e-09;
  Matches   30;  Conservative    9;  Mismatches   16;  Indels   2%;  Gaps    2;

Qy       6 DPGSHLVLGLTLAVGGAGFFYYTRTKKLMEDRpVQpYTNASSH 50
           ||| || | |  |||||||:::  |              | || |             : |
Db      49 DPGSRMVQTSpVTSSSIDTYSE-              SSCrGs             sKsn 93

Qy      51 GIVEQCCTSICSLYQLENYC 85
           |||:::|   | | :|| ||
Db      94 GIVDECCFQSCELRRLEMYC 133


Search completed: February  5, 2003, 12:47:29
Job time : 28.0502 secs
```

Run on: February 5, 2003, 12:41:22 ; Search time 1.78 Seconds
(without alignments)
290.219 Million cell updates/sec

Title: US-09-423-100-4
Perfect score: 463
Sequence: 1 EVNLHLGGSHLVEALYLV.....IVEQYISDQSLY.ENYCN 99

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2763 seqs, 1156252 residues

Total number of hits satisfying chosen parameters: 2763

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles2.pep:*

The Result No. is the number of results in order of decreasing score. The score is a measure of how well the query sequence is aligned with the database sequence. The Match % is the percent identity of the alignment. The length is the number of residues aligned. The Query Offset and DB Offset are the start of the alignment in the query and database sequence, respectively. The Description is the description of the database sequence. The E-value is the number of hits one can expect to see by chance when searching a database of this size with the query sequence, at the given score or better. The smaller the E-value, the more significant the match. The alignment length is the length of the alignment. The Match Count is the number of matches, conservative substitutions, insertions and deletions in the alignment. The Mismatch Count is the number of mismatches and conservative substitutions in the alignment. The Gap Count is the number of insertions and deletions in the alignment. The Gap Length is the total length of all gaps in the alignment.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	463	100.0	86	1	US-09-477-A-24 Sequence 24, Appli
2	463	100.0	96	2	US-09-154-836-4 Sequence 4, Appli
3	463	100.0	96	4	US-09-836-A-4 Sequence 4, Appli
4	463	100.0	97	1	US-08-476-A-4 Sequence 4, Appl
5	463	100.0	112	1	US-08-805-A-2 Sequence 12, Appl
6	463	100.0	113	1	US-08-781-A-2 Sequence 2, Appli
7	463	100.0	113	1	US-08-781-A-2 Sequence 2, Appli
8	463	100.0	113	1	US-08-785-A-2 Sequence 2, Appli
9	463	100.0	121	1	PCT-US95-05 Sequence 2, Appli
10	463	100.0	121	2	US-08-476-A-2 Sequence 15, Appl
11	463	100.0	121	2	US-08-476-A-2 Sequence 15, Appl
12	463	100.0	127	1	US-08-476-A-2 Sequence 8, Appl
13	463	100.0	127	1	US-08-781-A-2 Sequence 8, Appl
14	457	98.7	96	2	US-09-154-836-5 Sequence 5, Appli
15	457	98.7	96	4	US-09-836-A-5 Sequence 5, Appli
16	457	98.7	97	1	US-08-476-A-5 Sequence 5, Appl
17	456	98.5	1	1	US-08-476-A-1 Sequence 15, Appl
18	448	96.8	97	1	US-08-476-A-3 Sequence 3, Appl
19	444	95.9	97	1	US-08-476-A-6 Sequence 6, Appl
20	443	95.7	112	1	US-08-805-A-6 Sequence 6, Appli
21	443	95.7	113	1	US-08-781-A-6 Sequence 6, Appli
22	443	95.7	113	1	US-08-785-A-6 Sequence 6, Appli
23	438	94.6	97	1	US-08-476-A-7 Sequence 7, Appl
24	438	94.6	97	1	US-08-476-A-8 Sequence 8, Appl
25	438	94.6	97	1	US-08-476-A-9 Sequence 9, Appl
26	438	94.6	97	1	US-08-476-A-10 Sequence 10, Appl
27	438	94.6	97	1	US-08-476-A-11 Sequence 11, Appl
28	438	94.6	97	1	US-08-476-A-12 Sequence 12, Appl
29	438	94.6	97	1	US-08-476-A-13 Sequence 13, Appl
30	438	94.6	97	1	US-08-476-A-14 Sequence 14, Appl
31	438	94.6	97	1	US-08-476-A-15 Sequence 15, Appl
32	438	94.6	97	1	US-08-476-A-16 Sequence 16, Appl
33	438	94.6	97	1	US-08-476-A-17 Sequence 17, Appl
34	438	94.6	97	1	US-08-476-A-18 Sequence 18, Appl
35	438	94.6	97	1	US-08-476-A-19 Sequence 19, Appl
36	438	94.6	97	1	US-08-476-A-20 Sequence 20, Appl
37	438	94.6	97	1	US-08-476-A-21 Sequence 21, Appl
38	438	94.6	97	1	US-08-476-A-22 Sequence 22, Appl
39	438	94.6	97	1	US-08-476-A-23 Sequence 23, Appl
40	438	94.6	97	1	US-08-476-A-25 Sequence 25, Appl
41	438	94.6	97	1	US-08-476-A-26 Sequence 26, Appl
42	438	94.6	97	1	US-08-476-A-27 Sequence 27, Appl
43	438	94.6	97	1	US-08-476-A-28 Sequence 28, Appl
44	438	94.6	97	1	US-08-476-A-29 Sequence 29, Appl
45	438	94.6	97	1	US-08-476-A-30 Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-09-477-A-24
; Sequence 24, Application US/09477A
; Patent No. 6174861
; GENERAL INFORMATION:
; APPLICANT: DeLapeyre, Yves
; APPLICANT: Lawlor, David
; TITLE OF INVENTION: Use of Low Calcium
; FILE REFERENCE: 1-21-A
; CURRENT APPLICATION NUMBER: US/09/477,424
; CURRENT FILING DATE: 2000-01-04
; NUMBER OF SEQ IDS: 30
; SEQ ID NO 24
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-477-A-24

Query Match 100.0%; Score 463; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.2e-46;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 99
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 86

Qy 100 SL 99
|| ||
Db 87 SL 86

RESULT 2
US-09-154-836-4
; Sequence 4, Application US/09154836
; Patent No. 6174861
; GENERAL INFORMATION:
; APPLICANT: DeLapeyre, Yves
; APPLICANT: Lawlor, David
; TITLE OF INVENTION: Use of Low Calcium
; FILE REFERENCE: 1-21-A
; CURRENT APPLICATION NUMBER: US/09/154,836
; CURRENT FILING DATE: 1998-09-17
; NUMBER OF SEQ IDS: 30
; SEQ ID NO 4
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-154-836-4

Query Match 100.0%; Score 463; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 99
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 96

Qy 100 SL 99
|| ||
Db 97 SL 96

RESULT 3
US-09-836-A-4
; Sequence 4, Application US/09836A
; Patent No. 6174861
; GENERAL INFORMATION:
; APPLICANT: DeLapeyre, Yves
; APPLICANT: Lawlor, David
; TITLE OF INVENTION: Use of Low Calcium
; FILE REFERENCE: 1-21-A
; CURRENT APPLICATION NUMBER: US/09/836,A
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ IDS: 30
; SEQ ID NO 4
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-836-A-4

Query Match 100.0%; Score 463; DB 4; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.3e-46;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 99
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11 EVNLHLGGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 96

Qy 100 SL 99
|| ||
Db 97 SL 96

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,999
; FILING DATE:
; CLASSIFICATION: N
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonnell
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 021681.0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; BIOMATERIAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Xaa
; LOCATION: 1..9
; OTHER INFORMATION: /note= "Xaa = Any Amino Acid"

Query Match 100.0%; Score 45; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e-47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXXXXX 9
Db 1 XXXXXXXXX 9

RESULT 2
US-08-160-376A-4
; Sequence 4, Application US/08/160,376A
; Patent No. 5,670,345
; GENERAL INFORMATION:
; APPLICANT: Schneider, Francis
; APPLICANT: Gerl, Martin
; APPLICANT: Ludwig, Jurgen
; APPLICANT: Sabel, Walter
; TITLE OF INVENTION: Process for Purifying Human Serum Albumin
; TITLE OF INVENTION: Solutions Containing Stabilized Albumin
; TITLE OF INVENTION: Pasteurization Process
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,376A
; FILING DATE: 30-Nov-1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/971,444
; FILING DATE: 04-Nov-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonnell
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 021681.0008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; BIOMATERIAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Xaa
; LOCATION: 1..9
; OTHER INFORMATION: /note= "Xaa = Any Amino Acid"

Query Match 100.0%; Score 45; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e-47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXXXXXXX 9
Db 1 XXXXXXXXX 9

RESULT 3
US-08-160-376A-4
; Sequence 4, Application US/08/160,376A
; Patent No. 5,670,345
; GENERAL INFORMATION:
; APPLICANT: Schneider, Francis
; APPLICANT: Gerl, Martin
; APPLICANT: Ludwig, Jurgen
; APPLICANT: Sabel, Walter
; TITLE OF INVENTION: Process for Purifying Human Serum Albumin
; TITLE OF INVENTION: Solutions Containing Stabilized Albumin
; TITLE OF INVENTION: Pasteurization Process
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1500 K Street, N.W.
; CITY: Washington
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH, 1.44 MB STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 3.1
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,376A
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,444
; FILING DATE: 04-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara V. Mantel, Esq.
; REGISTRATION NUMBER: 31,387
; REFERENCE/DOCKET NUMBER: 08855/026 (PCT)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 253-8770
; TELEFAX: (914) 253-8259

; APPLICANT: Nov, Yelena Mikhailovna, Karl D.
; APPLICANT: Clark, Samuel A.
; APPLICANT: Holajan, Antonio E.
; APPLICANT: Quraishi, Christian
; APPLICANT: Krause, Fred
; APPLICANT: McGarry, Louis
; TITLE OF INVENTION: RECOMBINANT EXPRESSION IN INSECTS FROM
; TITLE OF INVENTION: SECRET BY CELL LINES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,582
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,427
; FILING DATE: 12-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,028
; FILING DATE: 09-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Hirshfeld, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 23
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-782-582-23

Query Match 100.0%; Score 460; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLQSGHEVEALYLVKPCVLYRDGYEPKFPESERMEGQVVEEGSGPDVSSQPLAEE 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 EVNQHLQSGHEVEALYLVKPCVLYRDGYEPKFPESERMEGQVVEEGSGPDVSSQPLAEE 84

Qy 61 SLQKGIVPAGGISLSLVQLENYGN 86
|||||||||||||||||||||||||
Db 85 SLQKGIVPAGGISLSLVQLENYGN 110

RESULT 9
; Sequence 23, Application US/08782582
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Proteolytic Peptide Compounds for Detection
; TITLE OF INVENTION: and Treatment Type I Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LARIVIERE, GRIMES
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1800
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20000
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/782,271
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,427
; FILING DATE: 12-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hirshfeld, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: UTSD:514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 23
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-782-271-23

Query Match 100.0%; Score 460; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 1.5e-47;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLQSGHEVEALYLVKPCVLYRDGYEPKFPESERMEGQVVEEGSGPDVSSQPLAEE 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 25 EVNQHLQSGHEVEALYLVKPCVLYRDGYEPKFPESERMEGQVVEEGSGPDVSSQPLAEE 84

Qy 61 SLQKGIVPAGGISLSLVQLENYGN 86
|||||||||||||||||||||||||
Db 85 SLQKGIVPAGGISLSLVQLENYGN 110

RESULT 9
PCT/US97/20000-2
; Sequence 23, Application PCT/US97/20000
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Proteolytic Peptide Compounds for Detection
; TITLE OF INVENTION: and Treatment Type I Diabetes
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LARIVIERE, GRIMES
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1800
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/20000
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```
qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 89
          ||||||||||||||||||||||||||||
db     1 SLQKRGIVEQCCTSICSLYQLENYCN 150


RESULT 1
US-09-423-100-4
; Sequence 4, Application US/09423100
; Patent No. 5821458
; GENERAL INFORMATION:
;   APPLICANT: ANDY, ROBIN L.
;   APPLICANT: LARSEN, ERIC R.
;   TITLE OF INVENTION: INSULIN SELECTIVE INSULIN ANALOGS
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: MORRISON & FOERSTER
;     STREET: 755 PAGE MILL ROAD
;     CITY: PALO ALTO
;     STATE: CALIFORNIA
;     COUNTRY: USA
;     ZIP: 94304-1018
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC Compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/423,100
;     FILING DATE: 12-NOV-1999
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BOZICEVIC, KARL
;     REGISTRATION NUMBER: 28,987
;     REFERENCE/DOCKET NUMBER: 21933.20274US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-477-7000
;     TELEFAX: (415)-434-4507
;     TELEX: 706141
;   INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 167 amino acids
;       TYPE: AMINO ACID
;       STRANDEDNESS:
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-423-100-4

  Query Match             100.0%;  Score 463;  DB 1;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-47;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 89
          ||||||||||||||||||||||||||||
Db     121 SLQKRGIVEQCCTSICSLYQLENYCN 149


RESULT 13
US-09-423-100-4
; Sequence 4, Application US/09423100
; Patent No. 5821458
; GENERAL INFORMATION:
;   APPLICANT: ANDY, ROBIN L.
;   APPLICANT: LARSEN, ERIC R.
;   TITLE OF INVENTION: INSULIN SELECTIVE INSULIN ANALOGS
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: MORRISON & FOERSTER
;     STREET: 755 PAGE MILL ROAD
;     CITY: PALO ALTO
;     STATE: CALIFORNIA
;     COUNTRY: USA
;     ZIP: 94304-1018
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC Compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/423,100
;     FILING DATE: 12-NOV-1999
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: BOZICEVIC, KARL
;     REGISTRATION NUMBER: 28,987
;     REFERENCE/DOCKET NUMBER: 21933.20274US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 415-477-7000
;     TELEFAX: (415)-434-4507
;     TELEX: 706141
;   INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 167 amino acids
;       TYPE: AMINO ACID
;       STRANDEDNESS:
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-09-423-100-4

  Query Match             100.0%;  Score 463;  DB 1;  Length 167;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-47;
  Matches   89;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEGS 60

Qy      61 SLQKRGIVEQCCTSICSLYQLENYCN 89
          ||||||||||||||||||||||||||||
Db     121 SLQKRGIVEQCCTSICSLYQLENYCN 149


RESULT 14
US-09-134-836-5
; Sequence 5, Application US/09134836
; Patent No. 5986048
; GENERAL INFORMATION:
;   APPLICANT: Bohnert, Franz-Josef
;   APPLICANT: Kellert, Reiko-Jo
;   TITLE OF INVENTION: Improved process for obtaining
;   TITLE OF INVENTION: insulin precursors having correctly bonded cystine bridges
;   NUMBER OF SEQUENCES: 7
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
;     ADDRESSEE: Dunner
;     STREET: 1300 I Street, N.W.
;     CITY: Washington
;     STATE: D.C.
;     COUNTRY: USA
;     ZIP: 20005-3315
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC Compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/09/134,836
;     FILING DATE:
;     CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Leslie McDonell
;     REGISTRATION NUMBER: 34,872
```

; REFERENCE/AGENT NUMBER: 0248.1600-0000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408 4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..99
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-423-100A-5

Query Match 98.7%; Score 497; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.6e-47;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 72

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 85
||||||||||||||||||||||||||
Db 73 SLQKRGIVEQCCTSICSLYQLENYCN 97

RESULT 5
US-09-423-100A-5
; Sequence 5, Application US/09423100A
; Patent No. 6380455
; GENERAL INFORMATION:
; APPLICANT: Rubroder, Franz-Josef
; Keller, Reinhold
; TITLE OF INVENTION: Improved process for obtaining
; insulin precursors having correctly bonded cystine bridges
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100A
; FILING DATE: 01-Apr-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/334,656
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonell
; REGISTRATION NUMBER: 34,872
; REFERENCE/DOCKET NUMBER: 0248.1600-0000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408 4400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..99
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-423-100A-5

Query Match 98.7%; Score 497; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.6e-47;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 13 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 72

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 85
||||||||||||||||||||||||||
Db 73 SLQKRGIVEQCCTSICSLYQLENYCN 97

Search completed: February 5, 2003, 12:47:13
Job time : 13.6438 secs

Title: US-09-423-100-4
Perfect score: 463

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

SUMMARIES

ALIGNMENTS

; APPLICANT: Dutcupilo, Yves
; APPLICANT: [illegible], Henry
; TITLE OF INVENTION: DE-[illegible] VARIANTS
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/10/028,411
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/09/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ IDS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-411-4

Query Match 100.0%; Score 463; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 [illegible] 60
[illegible]
Db 1 [illegible] 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
[illegible]
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

RESULT 4
US-10-028-411-4
; Sequence 4, Application US/10028411
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Tao, Zhuoyi Fu
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/028,411
; FILING DATE: 19-Dec-2001
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/477,924
; FILING DATE: 05-JAN-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers II, Frank C
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020167-000100US
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-028-411-4

Query Match 100.0%; Score 463; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 [illegible] 60
[illegible]
Db 1 [illegible] 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86
[illegible]
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

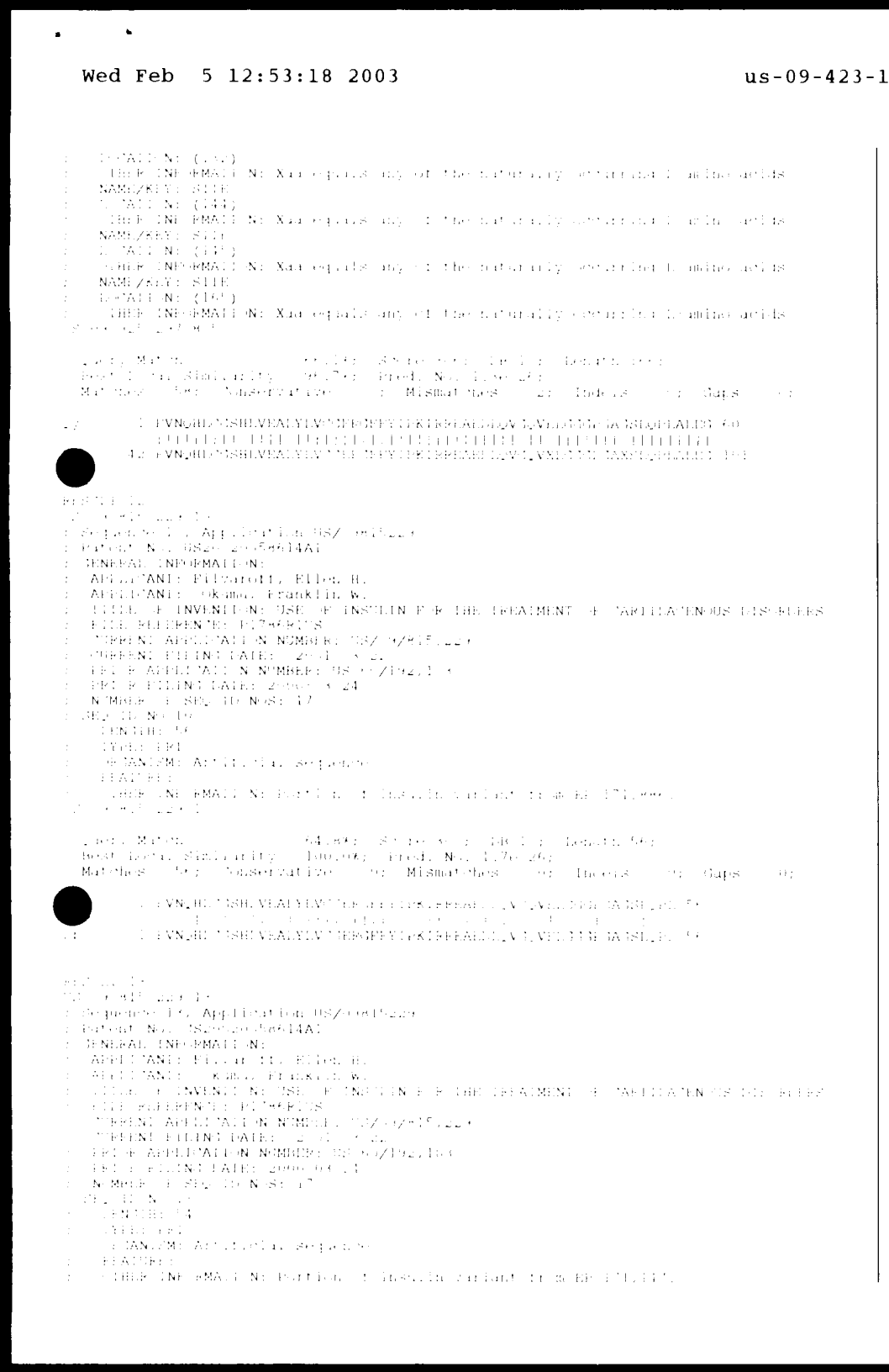
RESULT 4
US-10-028-411-4
; Sequence 4, Application US/10028411
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Rubrecht, Franz Josef
; Keller, Reinhold
; TITLE OF INVENTION: [illegible]
; [illegible]
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/947,563
; FILING DATE: 07-Sep-2001
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie Mcdonell
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..86
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-947-563-4

Query Match [illegible]; Score 463; DB 9; Length 86;
Best Local Similarity 100.0%; Pred. No. 4.3e-44;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 [illegible] 60
[illegible]
Db 1 [illegible] 60

Qy 61 SLQKRGIVEQCCTSICSLYQLENYCN 86

```
; Sequence 92, Application US/09423100A
; Patent No. US20010021519A1
; GENERAL INFORMATION:
;  APPLICANT: Sato, Seiji
;  APPLICANT: Hirashima, Naohiko
;  APPLICANT: Kodo, Toshiyuki
;  APPLICANT: Kondo, Masaaki
;  TITLE OF INVENTION: DNAS ENCODING NEW PROTEIN EXCITING AND PROCESSES FOR
;  TITLE OF INVENTION: PREPARING PURIFIED ERYTHROPOIETIN DERIVED FROM EXCISED N- T
;  TITLE OF INVENTION: DNAS
;  FILE REFERENCE: P02120
;  CURRENT APPLICATION NUMBER: US/09/423,100A
;  CURRENT FILING DATE: 1999-01-26
;  EARLIER APPLICATION NUMBER: JP10-07779/1998
;  EARLIER FILING DATE: 1998-03-24
;  NUMBER OF SEQ ID NOS: 96
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 92
;    LENGTH: 180
;    TYPE: PRT
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: Designated is
;    OTHER INFORMATION: an amino acid sequence of
;    OTHER INFORMATION: rhEPp MMPmpl0-(His)6-Linker-Met-Proinsulin.
;    US-09-423-100-4

  Query Match          100.0%;  Score 464;  DB 5;  Length 180;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-44;
  Matches   86;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       49 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 108

Qy       61 SLQKRGIVEQCCTSICSLYQLENYCN 86
            ||||||||||||||||||||||||||
Db      109 SLQKRGIVEQCCTSICSLYQLENYCN 134


RESULT 1
US-09-423-100-4
; Sequence 80, Application US/09925297
; Patent No. US20020006654A1
; GENERAL INFORMATION:
;  APPLICANT: Kabsch, Wolfgang et al.
;  TITLE OF INVENTION: Novel Actin, Its Coding Sequence and Antibodies
;  FILE REFERENCE: PA05
;  CURRENT APPLICATION NUMBER: US/09/925,297
;  CURRENT FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: EP01/05,749
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: US/225,279
;  PRIOR FILING DATE: 1999-08-13
;  NUMBER OF SEQ ID NOS: 928
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 80
;    LENGTH: 166
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: Site
;    LOCATION: (84)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    NAME/KEY: Site
;    LOCATION: (92)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    NAME/KEY: Site
;    LOCATION: (105)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    NAME/KEY: Site
;    LOCATION: (124)
;    OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;    NAME/KEY: Site
```



Multiple protein search using sw_m_bc

Run on: February 5, 2003, 12:21:21 ; Search time 1121.1 Seconds
(without alignments)
426.115 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 FVNQHLCGSHLVEALYLV[.....]VEQCCTSICSLYQLENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1044[illegible] seqs, 1[illegible] residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: [illegible]
2: [illegible]
3: [illegible]
4: [illegible]
5: [illegible]
6: [illegible]
7: [illegible]
8: [illegible]
9: [illegible]
10: [illegible]
11: [illegible]
12: [illegible]
13: [illegible]
14: [illegible]
15: [illegible]
16: [illegible]
17: [illegible]
18: [illegible]
19: [illegible]
20: [illegible]
21: [illegible]
22: [illegible]

Post-processing: Of the number of results predicted by chance to have a score greater than or equal to the score of the last hit printed, [illegible] is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	294	100.0	52	20	AAY42859	Human insulin prec
2	294	100.0	51	19	AAB29001	human pre-insulin
3	294	100.0	51	16	AAB72415	Bifunctin sequenc
4	294	100.0	51	19	AAB19500	Human pre-insulin
5	294	100.0	51	19	AAB19501	Human preinsulin
6	294	100.0	98	20	AAB78512	Fusion protein con
7	294	100.0	107	20	AAY12800	hGH mini-proinsuli
8	294	100.0	110	17	AAB28807	Sof proinsulin hyb
9	294	100.0	147	16	AAB71042	Marine factor alph
10	294	100.0	149	16	AAB71043	Marine factor alph
11	294	100.0	149	19	AAB19502	Human pre-insulin
12	294	100.0	149	19	AAB19503	Human pre-insulin
13	294	100.0	150	12	AAB71044	Marine factor alph
14	287	97.6	51	12	AAB71045	Insulin precursor
15	284.5	96.8	51	19	AAB29002	Human pre-insulin
16	284.5	96.8	51	20	AAB72414	Human pre-insulin
17	284.5	96.8	51	20	AAB72416	Bifunctin sequenc
18	284.5	96.8	51	20	AAB72417	Bifunctin sequenc
19	284.5	96.8	51	20	AAB72418	Bifunctin sequenc
20	278.5	94.7	51	21	AAB12771	Human pre-insulin
21	278.5	94.7	51	21	AAB12772	Human pre-insulin
22	278.5	94.7	51	22	AAM14214	Human pre-insulin
23	278.5	94.7	51	22	AAM14215	Human pre-insulin
24	278.5	94.7	51	22	AAM14216	Human pre-insulin
25	278.5	94.7	51	22	AAM14217	Human pre-insulin
26	278.5	94.7	51	22	AAM14218	Human pre-insulin
27	278.5	94.7	51	23	AAB12773	Human pre-insulin
28	278.5	94.7	51	23	AAB12774	Human pre-insulin
29	278.5	94.7	51	23	AAB12775	Human pre-insulin
30	278.5	94.7	51	23	AAB12776	Human pre-insulin
31	278.5	94.7	51	23	AAB12777	Human pre-insulin
32	278.5	94.7	51	23	AAB12778	Human pre-insulin
33	278.5	94.7	51	23	AAB12779	Human pre-insulin
34	278.5	94.7	51	23	AAB12780	Human pre-insulin
35	278.5	94.7	51	23	AAB12781	Human pre-insulin
36	278.5	94.7	51	23	AAB12782	Human pre-insulin
37	278.5	94.7	51	23	AAB12783	Human pre-insulin
38	278.5	94.7	51	23	AAB12784	Human pre-insulin
39	278.5	94.7	51	23	AAB12785	Human pre-insulin
40	278.5	94.7	51	23	AAB12786	Human pre-insulin
41	278.5	94.7	51	23	AAB12787	Human pre-insulin
42	278.5	94.7	51	23	AAB12788	Human pre-insulin
43	278.5	94.7	51	23	AAB12789	Human pre-insulin
44	278.5	94.7	51	23	AAB12790	Human pre-insulin
45	278.5	94.7	51	23	AAB12791	Human pre-insulin

ALIGNMENTS

RESULT 1
[illegible]
ID [illegible] standard; protein; [illegible] AA.
XX
AC [illegible];
XX
DT [illegible] (first entry)
XX
DE Human insulin precursor, SEQ ID [illegible].
XX
KW Insulin precursor; [illegible]; Major histocompatibility complex
KW [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
KW [illegible].
XX
OS Homo sapiens.
XX
PN WO9[illegible]-A1.
XX
PD [illegible]
XX
PF [illegible]
XX
PF [illegible]
XX
RA [illegible]
XX
RT [illegible]
XX
DR WPI; [illegible]
XX
PT New chimeric proteins [illegible] comprising [illegible] [illegible]
PT particularly [illegible] [illegible] [illegible].
XX

PS Claim 12; Page 29-30; 46pp; English.
XX
CC This sequence represents a human insulin precursor comprising insulin
CC A and B chains. This insulin precursor is a component of the chimeric
CC proteins hGH-mini-proinsulin (AAY42866) and the chimeric protein
CC qlcon in AAY42861. These chimeric proteins additionally contain an
CC N-terminal fragment of human growth hormone (hGH) and a cleavable
CC peptide linker (AAY42857). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor,
CC enabling it to fold correctly. The cleavable peptide linker has a
CC C-terminal Arg residue which enables the hGH portion of the
CC chimeric protein to be removed after folding has taken place. Production
CC of recombinant human insulin via an hGH-proinsulin chimeric protein can
CC provide human insulin with correctly linked cysteine bridges with
CC lower necessary procedural steps, and hence resulting in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intermolecular interactions among
CC the fusion proteins, thus allowing folding of the fused insulin precursor
CC at commercially useful high concentrations. The procedural steps of
CC cyanogen bromide cleavage, oxidative sulphitolysis and related
CC purification steps can thus be eliminated, along with the use of high
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 52 AA;

Query Match 100.0%; Score 294; DB 20; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.1e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

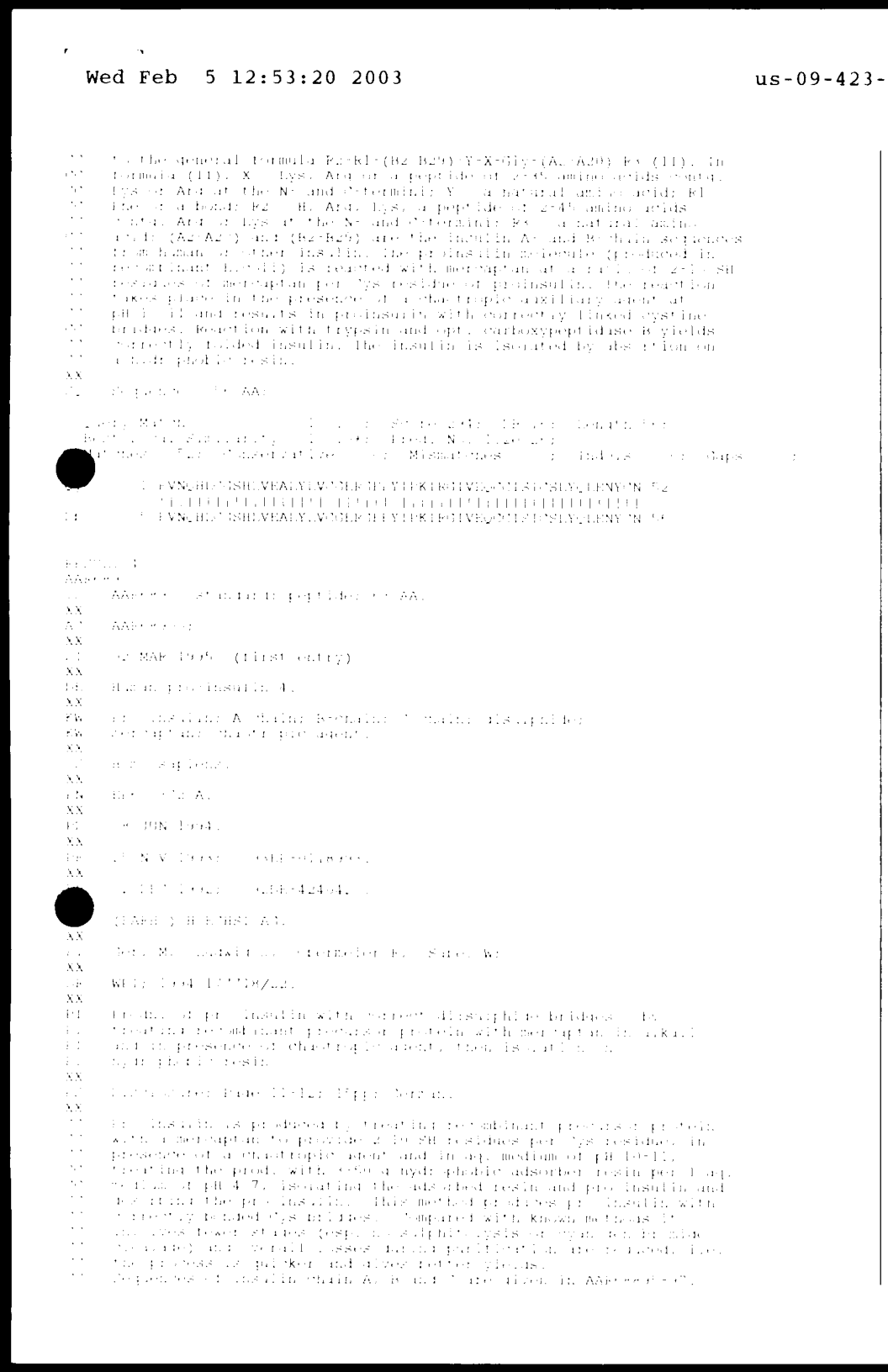
Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52

RESULT 2
AAR58990
ID AAR58990 standard; peptide; 56 AA.
XX
AC AAR58990;
XX
DT 02-MAR-1995 (first entry)
XX
DE Human pro-insulin 3.
XX
KW Pro-insulin; A chain; B chain; C chain; disulphide;
KW mercaptan; chaotropic agent.
XX
OS Homo sapiens.
XX
PN EP668292-A2.
XX
PD 23-AUG-1995.
XX
PF 09-FEB-1995; 95EP-0101748.
XX
PR 10-FEB-1994; 94DE-4404179.
XX
PA (BAYH) BAYER AG.
XX
PI Gerl M, Ludwig J, Obermeier R, Sabel W;
XX
DR WPI; 1995-253756/33.
XX
PT Prepn. of pro-insulin with correct disulphide bridges - by
PT treating recombinant precursor protein with mercaptan in alkali
PT and in presence of chaotropic agent, then isolation on
PT hydrophobic resin
XX
PS Disclosure; Page 12; 15pp; German.
XX
CC Pro-insulin is produced by treating recombinant precursor protein
CC with a mercaptan to provide 2 to 5H residues per Cys residue, in
CC presence of a chaotropic agent and in aq. medium of pH 10-11,
CC treating the prod. with 3-50 g hydrophobic adsorber resin per l aq.
CC medium of pH 4-7, isolating the adsorbed resin and pro-insulin, and
CC desorbing the pro-insulin. This method produces pro-insulin with
CC correctly bonded Cys bridges. Compared with known methods it
CC involves fewer stages (esp. no sulphitolysis or cyanogen bromide
CC cleavage) and overall losses during purification are reduced, i.e.
CC the process is quicker and gives better yields.
CC Sequences of insulin chain A, B and C are given in AAR60895-897.
CC Sequences of pro-insulin 1-4 are given in AAR60898-901.
XX
SQ Sequence 56 AA;

Query Match 100.0%; Score 294; DB 15; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.2e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
Db 5 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 56

RESULT 3
AAR78665
ID AAR78665 standard; protein; 56 AA.
XX
AC AAR78665;
XX
DT 03-APR-1996 (first entry)
XX
DE Proinsulin sequence 8.
XX
KW Proinsulin; post-translational modification; recombinant protein; a;
KW protein folding; conformation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..4
FT /label= B2
FT /note= "a peptide of 4 amino acids"
FT Peptide 5..34
FT /label= B1-(B2-B29)-Y
FT /note= "human insulin B chain"
FT Region 35
FT /label= X
FT Peptide 36..56
FT /label= GIY-(A2-A20)-FX
FT /note= "human insulin A chain"
XX
PN EP668292-A2.
XX
PD 23-AUG-1995.
XX
PF 09-FEB-1995; 95EP-0101748.
XX
PR 10-FEB-1994; 94DE-4404179.
XX
PA (BAYH) BAYER AG.
XX
PI Gerl M, Ludwig J, Obermeier R, Sabel W;
XX
DR WPI; 1995-253756/33.
XX
PT Isolation of insulin that is correctly post-translationally
PT processed - by reacting proinsulin with a mercaptan in the presence
PT of a chaotropic agent and partition, after absorption to hydrophobic
PT resin
XX
PS Example 2; Page 14; 16pp; German.
XX
CC The present sequence is an example of a proinsulin molecule correspl.



RESULT 1
AAB70012
ID AAB70012 standard; protein; 96 AA.
XX
AC AAB70012;
XX
DT 15-JAN-2000 (first entry)
XX
DE Basic protein conjg. proinsulin sequence 4.
XX
KW Proinsulin; post-translational modification; recombinant production;
KW protein folding; conformation.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 41..44
FT /label= K2
FT /note= "a peptide of 4 amino acids"
FT Peptide 45..74
FT /label= B1-B30 Y
FT /note= "human insulin B chain"
FT Bond 75
FT /label= X
FT Peptide 76..96
FT /label= chn (A2-A21) ks
FT /note= "human insulin A chain"
XX
PN WO9964462-A2.
XX
PD 16-DEC-1999.
XX
PF 10-JUN-1999; 99WO-EP004006.
XX
PR 11-JUN-1998; 98DE-1026179.
XX
PA (FARB) HOECHST AG.
XX
PI Frei M, Gerwig J, Kramer S, Sabel W;
XX
DR WPI; 2000-041754/36.
XX
PT Isolation of insulin that is correctly post-translationally
PT processed - by reacting proinsulin with a mercaptan in the presence
PT of a chaotropic agent and purifies, after absorption to hydrophobic
PT resin.
XX
PS Example 2; Page 40; 54pp; German.
XX
CC The present sequence is that of a fusion protein, produced in E.coli
CC which contains an example of a proinsulin molecule corresp.
CC to the general formula B2-K1 (B2-B29) Y X Gly-(A2-A20) K5 (I1) in
CC formula (II). X = Lys, Arg or a peptide of 2-35 amino acids contg.
CC Lys or Arg at the N- and C-termini; Y = a natural amino acid; K1
CC the or a bond; K2 = H, Arg, Lys, a peptide of 2-45 amino acids
CC contg. Arg or Lys at the N- and C-termini; K5 = a natural amino
CC acid; (A2-A20) and (B2-B29) are the insulin A- and B-chain sequences
CC of human or other insulin. The proinsulin molecule, released by
CC cyanogen bromide, is reacted with mercaptan at a ratio of 2-10 SH
CC positions of mercaptan per Cys residue of proinsulin. The reaction
CC takes place in the presence of a chaotropic auxiliary agent at
CC pH 10-11 and results in proinsulin with correctly linked cystine
CC bridges. Reaction with trypsin and opt. carboxypeptidase B yields
CC correctly folded insulin. The insulin is isolated by absorption on
CC a hydrophobic resin.
XX
SQ Sequence 96 AA;

Query Match 100.0%; Score 754; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.1e-20;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 96
||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 96

RESULT 2
AAY42866
ID AAY42866 standard; protein; 96 AA.
XX
AC AAY42866;
XX
DT 15-JAN-2000 (first entry)
XX
DE hGH with proinsulin C-peptide processing site.
XX
KW Insulin precursor; growth hormone fragment; fusion protein;
KW peptide; polymerase chain reaction; cleavable peptide;
KW production; yield.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9964462-A2.
XX
PD 16-DEC-1999.
XX
PF 10-JUN-1999; 99WO-EP004006.
XX
PR 11-JUN-1998; 98DE-1026179.
XX
PA (HOEC) HOECHST AG.
XX
PI Frei M, Gerwig J, Kramer S, Sabel W;
XX
DR WPI; 2000-041754/36.
XX
PT New chimeric proteins containing human growth hormone fragment, and
PT particularly for the production of human insulin
XX
PS Claim 14; Page 49; 54pp; German.
XX
CC This sequence represents a chimeric protein. hGH with proinsulin
CC This chimeric protein contains an N-terminal fragment of human growth
CC hormone (hGH) of the sequence given in AAY42865, a cleavable peptide
CC linker (AAY42867), and a human insulin precursor comprising insulin
CC A and B chains (AAY42869). The hGH portion of the chimeric protein acts
CC as an intramolecular chaperone (IMC) for the insulin precursor,
CC ensuring it is folded correctly. The cleavable peptide linker has a
CC C-terminal Arg residue which separates the hGH portion of the
CC chimeric protein to be removed after folding has taken place. The use
CC of recombinant human insulin (rhi) or hGH proinsulin chimeric proteins
CC provide human insulin with correctly linked cystine bridges with
CC fewer necessary procedural steps, and hence results in a higher yield
CC of human insulin. The IMC sequences not only protect insulin sequences
CC from intracellular degradation by a microorganism host, but also promote
CC the folding of the fused insulin precursor, facilitate the solubility of
CC the fusion protein and decrease the intracellular interaction between
CC the fusion proteins, thus allowing a higher of the fused insulin protein
CC at commercially useful cell concentrations. The proteins steps of
CC cyanogen bromide cleavage, cellular cell lysis and protein
CC purification steps can thus be eliminated, along with the use of high
CC concentrations of mercaptan or the use of hydrophobic absorbent resins.
XX
SQ Sequence 107 AA;

Query Match 100.0%; Score 754; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.1e-20;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 96
||||||||||||||||||||||||||||||||||||||||||||||||
Db 12 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 107

```
RESULT 8
AAR[illegible]
ID   AAR[illegible] standard; Protein; 116 AA.
XX
AC   AAR[illegible];
XX
DT   [illegible]-FEB-1997  (first entry)
XX
DE   S[illegible] proinsulin hybrid polypeptide.
XX
KW   insulin; proinsulin; hybrid polypeptide; protein folding;
KW   enzymatic cleavage; cyanogen bromide; sulphitolysis.
XX
OS   Homo sapiens.
XX
PN   W[illegible]20724 A1.
XX
PD   11-JUL-1996.
XX
PF   20-DEC-1994;   94WO-US13268.
XX
PR   2[illegible]-DEC-1994;   94WO-US13268.
PR   1[illegible]-JAN-1995;   95ZA-0000142.
XX
PA   (BIOT ) BIO-TECHNOLOGY GENERAL CORP.
XX
PI   Gorecki M,  Hartman JR,  Mendelovitz S;
XX
DR   WPI; 1996-[illegible]/[illegible].
DR   N-PSDB; AAT34670.
XX
PT   Recombinant insulin prodn. by correctly folding pro-insulin hybrid
PT   polypeptide - then enzymatic cleavage of folded product, does not
PT   require sulphite protection of SH nor use of cyanogen bromide
XX
PS   Example 1B; Figure 7; 69pp; English.
XX
CC   A new method for the production of recombinant human insulin
CC   comprises folding a hybrid polypeptide comprising proinsulin under
CC   conditions that permit correct disulphide bond formation and
CC   subjecting that folded protein to enzymatic cleavage.  The insulin
CC   produced can then be purified.  This sequence is a SOD insulin B
CC   chain Arg insulin A chain hybrid polypeptide and is encoded by
CC   the plasmid construct pDBAST-LAT.  Transformation of the proper
CC   E. coli host cells with pDBAST-LAT results in the efficient
CC   expression of the proinsulin hybrid polypeptide, useful for human
CC   insulin production. The method produces recombinant human insulin
CC   identical to the natural hormone.  Hazardous and cumbersome
CC   procedures involving cyanogen bromide and sulphitolysis to protect
CC   SH groups are avoided since the entire hybrid polypeptide folds
CC   efficiently to the native structure even with the leader attached
CC   and Cys unprotected.
XX
SQ   Sequence   116 AA;

  Query Match              100.0%;  Score 294;  DB 17;  Length 116;
  Best Local Similarity    100.0%;  Pred. No. 2.5e-26;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDLQVGQVELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 116
```

```
XX
KW   Human insulin precursor ArgB-1; diabetes; zinc ion complex;
KW   mating factor alpha 1.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..[illegible]
FT                   /label= mating factor alpha 1
FT   Peptide         [illegible]..1[illegible]
FT                   /label= B chain
FT   Peptide         1[illegible]..1[illegible]
FT                   /label= A chain
XX
PN   WO9507931-A.
XX
PD   23-MAR-1995.
XX
PF   16-SEP-1994;   94WO-DK00347.
XX
PR   17-SEP-1993;   93DK-0001041.
PR   02-FEB-1994;   94DK-0000029.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Andersen AS,  Halstrom JB,  Havelund S,  Jonassen I;
PI   Markussen J;
XX
DR   WPI; 1995-[illegible]/17.
DR   N-PSDB; AAQ86425.
XX
PT   Acylated insulin deriv. which may be present as a zinc ion
PT   complex - is used to treat diabetes and is rapid acting.
XX
PS   Example 5; Page 78; [illegible]pp; English.
XX
CC   AAQ86425 encodes AAR71692 mating factor alpha 1 insulin precursor
CC   ArgB-1. ArgB-1 comprises the B and A chains of a claimed human
CC   insulin derivative. In the final claimed compds, they are
CC   connected via disulphide bonds between Cys residues A7/B7 and
CC   A20/B19. The derivative, which may be present as a zinc ion
CC   complex, can be used as a fast acting treatment for diabetes.
XX
SQ   Sequence   1[illegible] AA;

  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible];  Pred. No. [illegible];
  Matches   52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN [illegible]

RESULT 10
AAR71694
ID   AAR71694 standard; Protein; 14[illegible] AA.
XX
AC   AAR71694;
XX
DT   2[illegible]-NOV-1995  (first entry)
XX
DE   Mating factor alpha 1 insulin precursor ArgB1, ArgB-1 N-term [illegible]
XX
KW   Human insulin precursor ArgB1, ArgB-1; diabetes; zinc ion complex;
KW   mating factor alpha 1; N-terminal EEAEAEAE.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Protein         1..8[illegible]
FT                   /label= mating factor alpha 1
FT   Peptide         8[illegible]..1[illegible]
```

```
RESULT 9
AAR71692
ID   AAR71692 standard; Protein; 137 AA.
XX
AC   AAR71692;
XX
DT   2[illegible]-NOV-1995  (first entry)
XX
DE   Mating factor alpha 1-insulin precursor ArgB-1.
```

```
FT                   /label= N-terminal peptide
FT   Peptide         <1..124
FT                   /label= B-chain
FT   Peptide         125..145
FT                   /label= A-chain
XX
PN   WO9[illegible]-A.
XX
PD   23-MAR-1995.
XX
PF   15-SEP-1994;    94WO-DK00347.
XX
PR   17-SEP-1993;    93DK-00104.
PR   02-FEB-1994;    94DK-00098.
XX
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Andersen AS;  Halstrom JB;  Bjerregaard S;  Jonasson I;
PI   Markussen J;
XX
DR   WPI; 1995-131814/17.
DR   N-PSDB; AAQ86429.
XX
PT   Acylated insulin deriv. which may be present as a zinc ion
PT   complex - is used to treat diabetes and is rapid acting.
XX
PS   Example 5; Page 85; 100pp; English.
XX
CC   AAQ86429 encodes AAK71694 mating factor alpha 1-insulin precursor
CC   Arg0b1, Arg0b31 N-terminal EEAEAEAER. The insulin precursor comprises
CC   the B and A chains of a claimed human insulin derivative preceded
CC   by the N-terminal amino acids EEAEAEAEr. In the final claimed compsn.
CC   they are covalently connected via disulphide bonds between Cys
CC   residues A7/B7 and A20/B19. The derivatives, which may be present
CC   as a zinc ion complex, can be used as a fast action treatment for
CC   diabetes.
XX
SQ   Sequence   145 AA;

  Query Match             100.0%;  Score 294;  DB 16;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-26;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 EVNQHLCGSHLVEALYLVCGERGFFYTPKTREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 145


RESULT 13
AAK71694
ID   AAK71694 standard; protein; 145 AA.
XX
AC   AAK71694;
XX
DT   19-JAN-2000  (first entry)
XX
DE   Mating factor alpha 1 insulin precursor Arg0b1, Arg0b31 N-terminal.
XX
KW   Human insulin precursor Arg0b1, Arg0b31; diabetes; zinc ion complex;
KW   mating factor alpha 1; N-terminal EEAEAEAER.
XX
OS   H.sapiens.
XX
FH   Key             Location/Qualifiers
FT   sig_peptide     1..89
FT                   /label= mating factor alpha1
FT   propeptide      89..94
FT                   /label= N-terminal peptide
FT   Peptide         95..125
FT                   /label= B-chain
FT   Peptide         126..146
FT                   /label= A-chain
XX
PN   WO9[illegible]-A.
XX
PD   23-MAR-1995.
XX
PF   15-SEP-1994;    94WO-DK00347.
XX
PR   17-SEP-1993;    93DK-00104.
PR   02-FEB-1994;    94DS-00098.
XX
PA   (NOVO ) NOVO NORDISK AS.
XX
PI   Andersen AS;  Halstrom JB;  Bjerregaard S;  Jonasson I;
PI   Markussen J;
XX
DR   WPI; 1995-131814/17.
DR   N-PSDB; AAQ86429.
XX
PT   Acylated insulin deriv. which may be present as a zinc ion
PT   complex - is used to treat diabetes and is rapid acting.
XX
PS   Example 5; Page 85; 100pp; English.
XX
CC   AAQ86429 encodes AAK71694 mating factor alpha 1-insulin precursor
CC   Arg0b1, Arg0b31 N-terminal EEAEAEAER. The insulin precursor comprises
CC   the B and A chains of a claimed human insulin derivative preceded
CC   by the N-terminal amino acids EEAEAEAEr. In the final claimed compsn.
CC   they are covalently connected via disulphide bonds between Cys
CC   residues A7/B7 and A20/B19. The derivatives, which may be present
CC   as a zinc ion complex, can be used as a fast action treatment for
CC   diabetes.
XX
SQ   Sequence   145 AA;

  Query Match             100.0%;  Score 294;  DB 16;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-26;
  Matches  52;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 52
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 EVNQHLCGSHLVEALYLVCGERGFFYTPKTREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 145


RESULT 14
AAY42[illegible]
ID   AAY42[illegible] standard; protein; [illegible] AA.
XX
AC   AAY42[illegible];
XX
DT   19-JAN-2000  (first entry)
XX
DE   Chimeric protein, SEQ ID 7.
XX
KW   Insulin precursor; insulin analogue; therapeutic; intranasal;
KW   soluble; controlled release; chimeric protein; recombinant DNA;
KW   production; yield.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9950302-A1.
XX
PD   07-[illegible]-1999.
XX
PF   31-MAR-1999;    99WO-EP02179.
XX
PR   31-MAR-1998;    98WO-EP01918.
XX
PA   (ZHONG ) ENCHUA GANJIE'H RES ECHNOLOGY CO.
XX
PI   Gan Y;
XX
DR   WPI; [illegible]-[illegible]/[illegible].
XX
```

PT New chimeric proteins containing human growth hormone fragment, used
PT particularly for the production of human insulin.
XX
PA [illegible] [illegible] Pty Limited.
XX
CC This sequence represents a chimeric protein, which contains an
CC N-terminal fragment of human growth hormone (hGH) of the sequence given
CC in AAY42e56, a cleavable peptide linker (AAY42e57), and a human insulin
CC precursor B-A sequence (B chain A acid sequence (AAY42e58). The hGH portion
CC of the chimeric protein is an intermediate shape to (IM') for
CC the insulin precursor, such that it is less susceptible to chimeric
CC peptide linker has a C-terminal Arg residue which enables the hGH portion
CC of the chimeric protein to be removed after folding has taken place.
CC Production of recombinant human insulin via an hGH-proinsulin chimeric
CC protein can provide human insulin with correctly linked cysteine bridges
CC with fewer necessary procedural steps, and hence result in a higher
CC yield of human insulin. The IM' sequences not only protect insulin
CC sequences from intracellular degradation by a microorganism host, but
CC also promote the folding of the fused insulin precursors, facilitate the
CC stability of the fusion protein and decrease the internal
CC interactions among the fusion proteins, thus allowing folding of the
CC fused insulin precursor at commercially useful high concentrations. The
CC proteolytical steps of cyanogen bromide cleavage, oxidative sulphitolysis
CC and related purification steps can thus be eliminated, along with the use
CC of toxic chemicals such as mercaptan. The use of hydrophobic adsorbent
CC resins.
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible]; DB [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; protein; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible] (first entry)
XX
DE [illegible] chain chain with a Lys residue linking the A and B chains.
XX
KW Insulin fusion protein; pro insulin analogue; tandem[illegible];
KW Lys-Lys bridges.
XX
FH Key Location/qualifiers
FT misc_difference 36
FT /label Lys residue linking insulin B chain to A chain
FT sig_peptide [illegible]
FT /label Insulin B chain
FT sig_peptide [illegible]
FT /label Insulin A chain
XX
SQ Sequence [illegible] AA;
XX
CC [illegible] MAY [illegible]
XX
Qy [illegible]
Db [illegible]
XX
RA (Arch) H E'BRIEN A;
XX
RL K Iver Kr. Bloss G. Uhlmann E. Wallmeier H;
XX
DR WPI; [illegible]/[illegible].
DR N-PSDB; [illegible].
XX
PT New [illegible] [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
PT [illegible]
XX
PS Disclosure; [illegible] Pages.
XX
CC [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
CC comprises [illegible] [illegible] [illegible] [illegible] [illegible]
CC may be [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
CC gene was prepared by the phosphotriester method.
CC See also AA[illegible].
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible]; DB [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT [illegible]
AAB[illegible]
ID AAB[illegible] standard; protein; [illegible] AA.
XX
AC AAB[illegible];
XX
DT [illegible] (first entry)
XX
DE Example of human insulin precursor.
XX
KW Human insulin; B-chain; A-chain; [illegible].
XX
OS Homo sapiens.
XX
RN [illegible] A.
XX
RC [illegible] MAY [illegible]
XX
RP [illegible] MAY [illegible]
RP [illegible] MAY [illegible]
RP [illegible] [illegible]
XX
RA (N.V.) N.V. N.V. [illegible] A/S;
XX
RT Manufacture [illegible] [illegible] [illegible] [illegible] [illegible]
RT Novo Nordisk A/S [illegible]
XX
DR WPI; [illegible]/[illegible].
XX
PT Human insulin precursor [illegible] [illegible] [illegible] [illegible]
PT [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
XX
PS Claim 1; page [illegible] [illegible].
XX
CC This human insulin precursor [illegible] [illegible] [illegible]
CC [illegible] between the A and B chains and [illegible] [illegible]
CC [illegible] [illegible] [illegible] [illegible] [illegible] [illegible]
CC Strains transformed with DNA encoding this precursor can be
CC cultured to secrete it in high yields. The precursor can be
CC converted into [illegible] human insulin. [illegible]
CC See also AA[illegible].
XX
SQ Sequence [illegible] AA;

Query Match [illegible] Score [illegible]; DB [illegible] Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];

Matches 51; Conservative 32; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 82

Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKSKREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 82

RESULT 19
AAB67###
ID AAB67### standard; protein; ## AA.
XX
AC AAB67###;
XX
DT ##-###-#### (first entry)
XX
DE ## Ara (Bos ##) Human insulin amorphous, monospherical deriv.
XX
KW B## ##; insulin; recombinant production; amorphous; monospherical; ##ms;
KW diabetes mellitus.
XX
OS Bos sapiens (produced recombinantly in Escherichia coli).
XX
FH Key Location/Qualifiers
FT ##_###
FT /date= ##-###-####
FT Protein ##..##
FT /label= insulin_B-chain
XX
FT ##_###
FT /label= insulin_A-chain
XX
PN EP###### A.
XX
PD ## NOV ####.
XX
PF ## APR ####; ##EP-######.
XX
PR ## APR ####; ##EE-######.
XX
PA (FARB) HOECHST AG.
XX
PI Bell E, Nelsen K, Obermeier R, Sabel W;
XX
DR WPI; ####-######/##.
XX
CC Appl. as mono spherical form of insulin deriv., ### for ###
CC diabetes mellitus, are prepared by dissolving in
CC isopropanol, are stable when dried or in suspension
XX
SQ Example 2; Page 5; 1 pp; German.
XX
CC This sequence is a specific example of an insulin derivative which
CC can be obtained in amorphous, monospherical form by dissolving in
CC an n-propanol/buffer mixture (pH 4.5-6.5) having n-propanol content
CC ##% relative to water. The solution is then diluted with water to
CC ###.5% n-propanol content to below ##%. The resulting insulin
CC preparation is stable and can be used for the treatment of diabetes
CC mellitus.
XX
SQ Sequence ## AA;

Query Match ##.##%; Score ###.#; DB #; Length ##;
Best Local Similarity ##.#%; Pred. No. #.##e-##;
Matches ##; Conservative ##; Mismatches #; Indels #; Gaps #;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT RREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 82
||| ||||||||||||||||||:|:|| ||||||||| ||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 82

Search completed: ## ###, ####, 12:44:28
CPU time: ##.#### secs

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 FVN_HLGSHLVEALYLV.....VE_GTSERGFF_IENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 286228 seqs, 90144422 residues

Total number of hits satisfying chosen parameters: 286228

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Prob. Nn is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1	294.0	100.0	51	1	INWHIP	insulin - sperm wh
2	278.5	94.7	51	1	INWHP	insulin - finback
3	278.5	94.7	51	1	INHEL	insulin - elephant
4	277	94.2	96	2	P27782	epidermal growth f
5	276.5	94.0	51	1	INHY	insulin - hamster
6	275.5	93.7	51	1	INMSSI	insulin - Brazilia
7	275.5	93.7	51	2	A25251	insulin precursor
8	267	90.8	110	1	IRHU	insulin precursor
9	267	90.8	110	1	B42179	insulin precursor
10	267	90.8	110	2	A42179	insulin precursor
11	267	90.8	110	2	JQ0178	insulin precursor
12	265.5	90.3	51	1	INWHIS	insulin - sei whal
13	265.5	90.3	51	1	INGI	insulin - goat
14	265.5	90.3	51	2	INOMA	insulin - Arabian
15	263	89.5	84	1	IRFI	insulin precursor
16	263	89.5	110	1	INRB	insulin precursor
17	263	89.5	51	1	IN*L	insulin - rat
18	263	89.5	110	1	IRGI	insulin precursor
19	262.5	89.3	51	1	INMKS	insulin - common s
20	261	88.8	110	2	I48619	insulin precursor
21	260.5	88.6	105	1	IRPO	insulin precursor
22	259.5	88.2	51	2	A26962	insulin - North Am
23	259.5	88.2	90	1	INSH	insulin precursor
24	259	88.1	86	1	IRBE	insulin precursor
25	258.5	87.9	51	1	IN*G	insulin - Chinchi
26	256	87.1	109	2	A42995	insulin precursor
27	256	87.1	109	1	INMS1	insulin 1 precurso
28	254	86.4	109	1	IRRT	insulin 1 precurso
29	254.5	86.6	51	1	IN1S	insulin - goose
30	254	86.4	51	1	IN*M	insulin - mouse
31	253.5	86.2	51	1	IN*R	insulin - rat
32	253.5	86.2	51	1	IN*S	insulin - sheep
33	253.5	86.2	51	1	IN*Y	insulin - yak
34	253.5	86.2	51	1	IN*Z	insulin - Chinese
35	253.5	86.2	51	1	IN*BA	insulin - bactrian
36	253.5	86.2	51	1	IN*CA	insulin - camel
37	253.5	86.2	51	1	IN*DA	insulin - dromedar
38	253.5	86.2	51	1	IN*EL	insulin - elephant
39	253.5	86.2	51	1	IN*GI	insulin - goat
40	253.5	86.2	51	1	IN*HA	insulin - Arabian
41	253.5	86.2	51	1	IN*HO	insulin - horse
42	253.5	86.2	51	1	IN*LA	insulin - llama
43	253.5	86.2	51	1	IN*PA	insulin - pacific
44	253.5	86.2	51	1	IN*SA	insulin - saiga
45	253	86.1	51	1	IN*XA	insulin - Asian

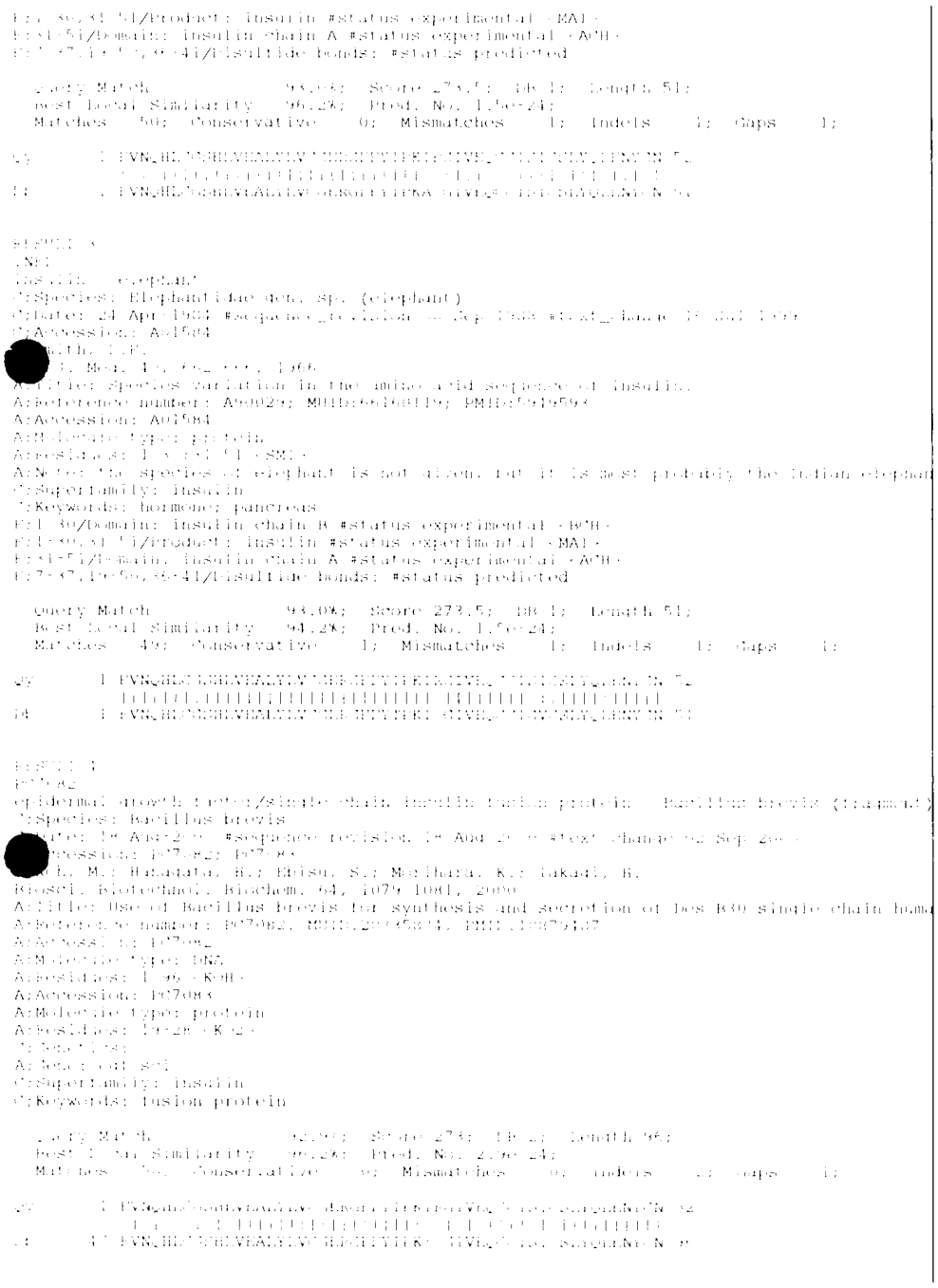
ALIGNMENTS

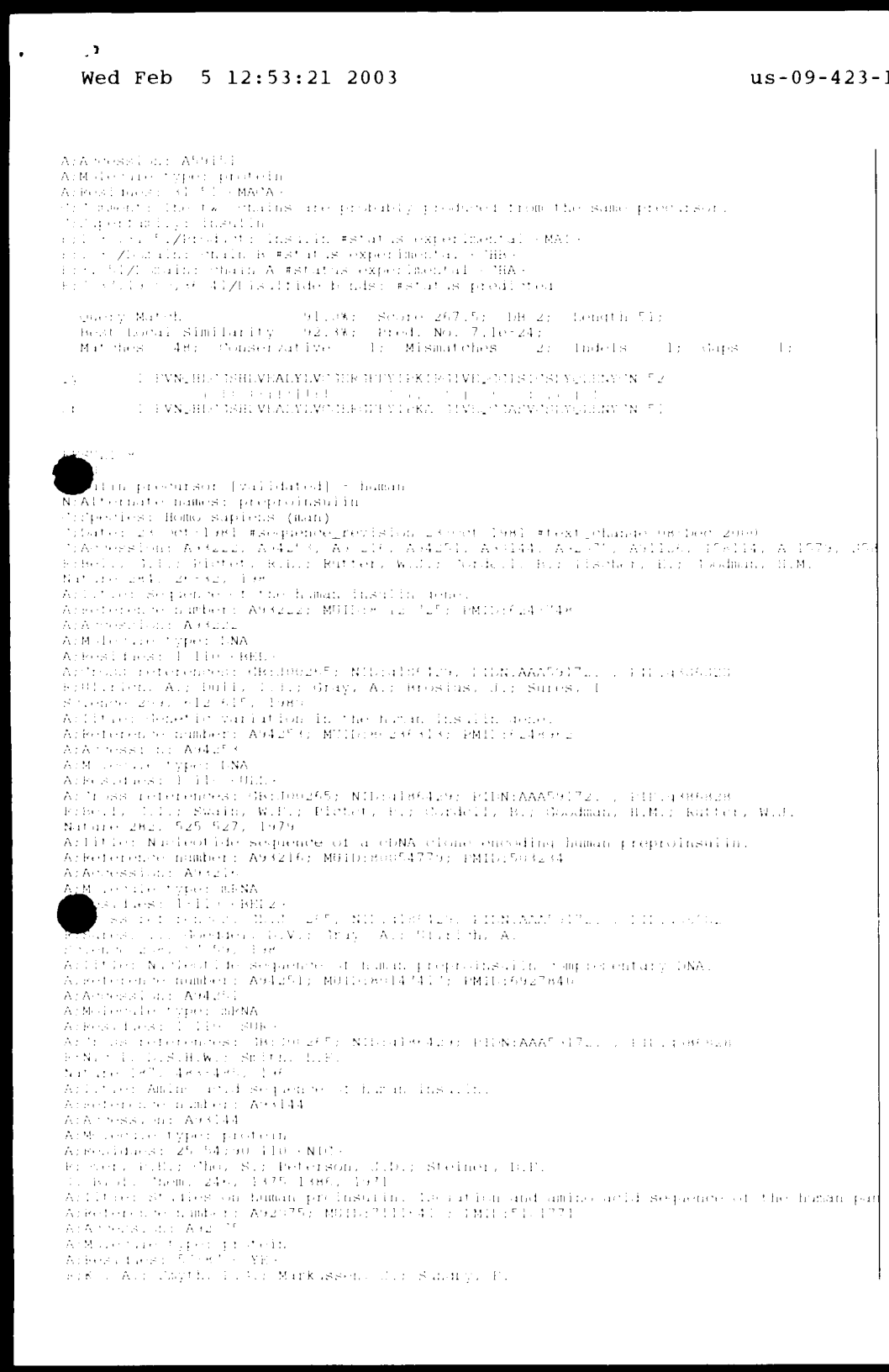
RESULT 1
INWHIP
insulin - sperm whale
C;Species: Eukaryota; Chordata; Mammalia; Cetacea; Odontoceti; Physeteridae; Physeter
C;Date: 31-Mar-1982 #sequence_revision 31-Mar-1982 #text_change 16-Jul-1987
C;Accession: A01241; A00006
R;Ishihara, Y.; Saito, H.; Doi, Y.; Fujino, M.
Nature 181, 1498-1499, 1958
A;Title: Structure of sperm whale insulin and the evolution of insulin
A;Reference number: A01241
A;Accession: A01241
A;Molecule type: protein
A;Residues: 1-51
R;Harris, J.I.; Sanger, F.; Naughton, M.A.
Arch. Biochem. Biophys. 77, 1-22, 1958
A;Title: Species differences in insulin
A;Reference number: A00006
A;Accession: A00006
A;Molecule type: protein
A;Residues: 1-51
C;Superfamily: Insulin
C;Keywords: hormone; pancreas
F;1-21/Domain: Insulin B chain #status experimental
F;1,22/Disulfide bonds: Insulin #status experimental
F;7-37,29-51,36-41/Disulfide bonds: #status predicted

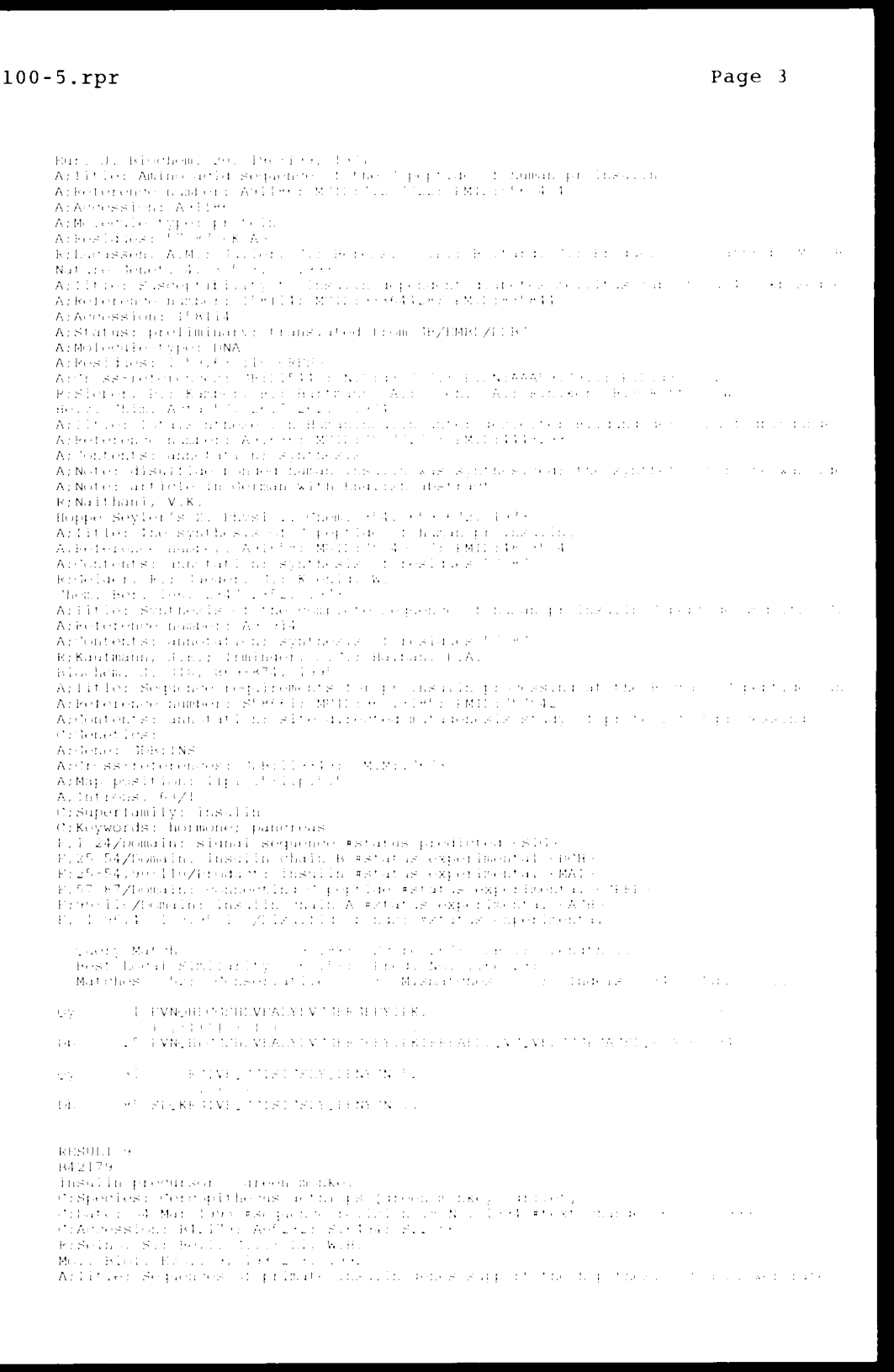
Query Match 100.0%; Score 294; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA 52
||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA 52

RESULT 2
INWHIP
Insulin - finback whale (fragment)
C;Species: Balaenoptera physalus (finback whale)
C;Date: 31-Mar-1982 #sequence_revision 31-Mar-1982 #text_change 16-Jul-1987
C;Accession: A01242
Kitama, H.; Ikemi, K.; Kaneko, S.; Nakata, E.
J. Biochem. 96, 287-293, 1984
A;Title: The amino acid sequence of fin whale insulin.
A;Reference number: A01242
A;Accession: A01242
A;Molecule type: protein
A;Residues: 1-51
C;Superfamily: Insulin
C;Keywords: hormone; pancreas
F;1-21/Domain: Insulin B chain #status experimental







```
A;Reference number: A42179; MUID:92219953; PMID:1560757
A;Accession: B42179
A;Molecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61092; NID:g22803; PIDN:CAA43405.1; PID:g22805
A;Note: sequence extracted from NCBI backbone (NCBIN:95185, NCBIP:95194)
R;Peterson, J.D.; Nehrlich, S.; Oyer, P.E.; Steiner, D.F.
J. Biol. Chem. 247, 4866-4871, 1972
A;Title: Determination of the amino acid sequence of the monkey, sheep, and dog proinsul
A;Reference number: A92111; MUID:72258016; PMID:4626369
A;Accession: A05232
A;Molecule type: protein
A;Residues: 57-87 <PEI>
C;Genetics:
A;Introns: 63/1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status predicted <BCH>
  5-54,90-110/Product: insulin #status predicted <MAT>
  7-87/Domain: connecting peptide #status experimental <CHA>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status predicted

  Query Match             90.8%;  Score 267;  DB 2;  Length 110;
  Best Local Similarity   60.5%;  Pred. No. 1.6e-23;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels   34;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT                               30
              ||||||||||||||||||||||||||||||
Db         25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy         31 ----RGIVEQCCTSICSLYQLENYCN 52
                  ||||||||||||||||||||||
Db         85 SLQKRGIVEQCCTSICSLYQLENYCN 110


RESULT 10
A42179
insulin precursor - chimpanzee
C;Species: Pan troglodytes (chimpanzee)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C;Accession: A42179; S22058
R;Seino, S.; Bell, G.I.; Li, W.H.
Mol. Biol. Evol. 9, 193-203, 1992
A;Title: Sequences of primate insulin genes support the hypothesis of a slower rate of m
A;Reference number: A42179; MUID:92219953; PMID:1560757
 Accession: A42179
  atus: preliminary
  olecule type: DNA
A;Residues: 1-110 <SEI>
A;Cross-references: EMBL:X61089; NID:g38251; PIDN:CAA43403.1; PID:g38252
A;Note: sequence extracted from NCBI backbone (NCBIP:95667)
C;Genetics:
A;Introns: 63/1
C;Superfamily: insulin

  Query Match             90.8%;  Score 267;  DB 2;  Length 110;
  Best Local Similarity   60.5%;  Pred. No. 1.6e-23;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels   34;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT------                         30
              ||||||||||||||||||||||||||||||
Db         25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy         31 ----RGIVEQCCTSICSLYQLENYCN 52
                  ||||||||||||||||||||||
Db         85 SLQKRGIVEQCCTSICSLYQLENYCN 110


RESULT 11
JQ0178
insulin precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 16-Jul-1999
C;Accession: JQ0178
R;Wetekam, W.; Groneberg, J.; Leineweber, M.; Wengenmayer, F.; Winnacker, E.L.
Gene 19, 179-183, 1982
A;Title: The nucleotide sequence of cDNA coding for preproinsulin from the primate Ma
A;Reference number: JQ0178; MUID:83080474; PMID:6184262
A;Accession: JQ0178
A;Molecule type: mRNA
A;Residues: 1-110 <WET>
A;Cross-references: GB:J00336; NID:g342121; PIDN:AAA36849.1; PID:g342122
C;Superfamily: insulin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54,90-110/Product: insulin #status predicted <MAT>
F;25-54/Domain: insulin chain B #status predicted <BCH>
F;55-89/Domain: insulin connecting C peptide #status predicted <CPE>
F;90-110/Domain: insulin chain A #status predicted <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status predicted

  Query Match             90.8%;  Score 267;  DB 2;  Length 110;
  Best Local Similarity   60.5%;  Pred. No. 1.6e-23;
  Matches   52;  Conservative    0;  Mismatches    0;  Indels   34;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT------------------------------ 30
              ||||||||||||||||||||||||||||||
Db         25 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 84

Qy         31 ----RGIVEQCCTSICSLYQLENYCN 52
                  ||||||||||||||||||||||
Db         85 SLQKRGIVEQCCTSICSLYQLENYCN 110


RESULT 12
INWHIS
insulin - sei whale
C;Species: Balaenoptera borealis (sei whale)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C;Accession: A01582
R;Ishihara, Y.; Saito, T.; Ito, Y.; Fujino, M.
Nature 181, 1468-1469, 1958
A;Title: Structure of sperm- and sei-whale insulins and their breakdown by whale peps
A;Reference number: A01582
A;Accession: A01582
A;Molecule type: protein
A;Residues: 1-30;31-51 <ISH>
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-30/Domain: insulin chain B #status experimental <BCH>
F;1-30,31-51/Product: insulin #status experimental <MAT>
F;31-51/Domain: insulin chain A #status experimental <ACH>
F;7-37,19-50,36-41/Disulfide bonds: #status predicted

  Query Match             89.6%;  Score 263.5;  DB 1;  Length 51;
  Best Local Similarity   92.3%;  Pred. No. 2e-23;
  Matches   48;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
              ||||||||||||||||||||||||||||  |||||| | ||||||||||
Db          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA-GIVEQCCASTCSLYQLENYCN 51


RESULT 13
INGT
insulin - goat
C;Species: Capra aegagrus hircus (domestic goat)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 16-Jul-1999
C;Accession: A01586
R;Smith, L.F.
Am. J. Med. 40, 662-666, 1966
A;Title: Species variation in the amino acid sequence of insulin.
A;Reference number: A90029; MUID:66160119; PMID:5949593
A;Accession: A01586
A;Molecule type: protein
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

SW protein protein search, using sw model

Run on: February 5, 2003, 12:42:21 ; Search time 4.05 Seconds
(without alignments)
462.278 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 FVNQHLCGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476528 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	275.5	93.0	51	1	INS_BALPH	P01312 balaenopter
2	272.5	92.0	51	1	INS_PHYMA	P01315 physeter ma
3	268.5	91.3	51	1	INS_ACOCA	P01324 acomys cahi
4	267	90.8	110	1	INS_CERAE	P30407 cercopither
5	267	90.8	110	1	INS_HUMAN	P01308 homo sapien
6	267	90.8	110	1	INS_MACFA	P80405 macaca fasc
7	267	90.8	110	1	INS_PANTR	P80416 pan troglod
8	264.5	89.6	51	1	INS_BALMU	P01314 balaenopter
9	264.5	89.6	51	1	INS_CAMDR	P01320 camelus dro
10	264.5	89.6	51	1	INS_CAPHI	P01319 capra hircu
11	264	89.5	108	1	INS_PIG	P01315 sus scrofa
12	264	89.5	110	1	INS_RABIT	P01311 oryctolagus
13	262.5	89.0	51	1	INS_BOSMU	P01318 bubalus bub
14	262	89.1	110	1	INS_CANFA	P01321 canis famil
15	260	88.4	110	1	INS_CRIGR	P01314 cricetulus
16	258.5	87.9	108	1	INS_BOVIN	P01317 bos taurus
17	257	87.4	110	1	INS_DASAP	Q62597 pseudomys n
18	256.5	87.2	51	1	INS_DIDMA	P18168 didelphis m
19	254.5	86.6	105	1	INS_SHEEP	P01318 ovis aries
20	252	85.7	86	1	INS_EQUCB	P01319 equus caball
21	251.5	85.5	51	1	INS_CHIBR	P01327 chinchilla
22	251	85.4	108	1	INS_AOTTR	P10604 aotus trivi
23	250	85.0	109	1	INS1_MOUSE	P01325 mus musculu
24	249	84.7	110	1	INS1_RAT	P01322 rattus norv
25	248.5	84.5	51	1	INS_ANSAN	P07454 anser anser
26	248	84.4	110	1	INS2_MOUSE	P01326 mus musculu
27	247	84.4	110	1	INS2_RAT	P01323 rattus norv
28	246	83.7	52	1	INS_ACIGU	P01428 acipenser g
29	244.5	83.2	51	1	INS_HYSCR	P01328 hystrix cri
30	244.5	83.2	51	1	INS_TRIAM	P35557 triakis sem
31	243.5	82.9	104	1	INS_SELRO	P01463 scrisphorus
32	243.5	82.9	51	1	INS_ORNAN	Q07977 ornithorhyn
33	243.5	82.9	107	1	INS_CHICK	P01332 gallus gall
34	243.5	82.9	51	1	INS_ANAPL	P01329 anas platyr
35	243.5	82.9	51	1	INS_ALLMI	P22466 alligator m
36	241	82.0	52	1	INS_CERSI	P01451 cervus sika
37	239.5	81.5	51	1	INS_CAVPO	P01326 cavia porcu
38	237.5	80.8	51	1	INS_ORYAF	P16044 oryzias lat
39	236.5	80.4	104	1	INS1_XENLA	P01327 xenopus lae
40	236.5	80.4	104	1	INS2_XENLA	P01328 xenopus lae
41	232	78.9	52	1	INS_PETMA	P14966 petromyzon
42	216	73.5	51	1	INS_ETAPE	P09467 pterois vol
43	213	72.4	52	1	INS_AMIPU	P20657 amia calva
44	211	71.8	51	1	INS_NOTKI	P01347 notothenia
45	211	71.8	51	1	INS2_CHUJU	P17691 hydrolagus

ALIGNMENTS

RESULT 1
INS_BALPH
ID INS_BALPH STANDARD; PRT; 51 AA.
AC P01312;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Insulin.
GN INS.
OS Balaenoptera physalus (Fin whale) (Common rorqual).
OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
OX NCBI_TaxID=9770, 9755;
RN [1]
RP PARTIAL SEQUENCE.
RC SPECIES=Physalus;
RA Nama H., Itoh N., Sakaki S., Narita K.;
RT "The amino acid sequence in fin whale insulin.";
RL J. Biochem. 96:1285-1287(1984).
RN [2]
RP SEQUENCE.
RC SPECIES=Catodon;
RA Ishihara T., Shii K., Ito Y., Baba H.;
RT "Structure of sperm whale and fin whale insulins and their interaction
RT whale pepsin.";
RL Nature 161:1469 1470(1948).
RN [3]
RP SEQUENCE.
RC SPECIES=Catodon;
RA Harris J.I., Sanger F., Naughton M.A.;
RT "Species differences in insulin.";
RL Arch. Biochem. Biophys. 65:427 438(1956).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES THE UPTAKE OF GLUCOSE BY MANY TISSUES. ALSO AFFECTS
CC FATTY ACIDS, KETONE BODIES, GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HEXAMER (1 A CHAIN AND 2 B CHAINS). ALSO
CC DIMERS AND MONOMERS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR PIR; A01018; INWHF.
DR PIR; A93413; INWHF.
DR HSSP; P01317; 1ASH.
DR InterPro; IPR003235; Insulin.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGc; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Signal; Disulfide bond.
FT CHAIN 1 21 INSULIN B CHAIN.
FT PEPTIDE 22 30 C-PEPTIDE.
FT CHAIN 31 51 INSULIN A CHAIN.
FT DISULFID 7 31 INTERCHAIN.
FT DISULFID 19 24 INTRACHAIN.

```
ID   [illegible]      xx      41
SQ   SEQUENCE   51 AA;  5776 MW;  [illegible]
  Query Match             [illegible]%;  Score 274.5;  DB 1;  Length 51;
  Best Local Similarity   96.2%;  Pred. No. 2.2e-26;
  Matches   50;  Conservative    0;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 50
           | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVEQCCASVCSLYQLENYCN 51


RESULT 2
INS_ELEMA
ID   INS_ELEMA      STANDARD;      PRT;    51 AA.
AC   P01345;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   02-FEB-1996 (Rel. 33, Last annotation update)
DE   Insulin.
GN   INS.
OS   Loxodonta africana (African elephant).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Proboscidea; Elephantidae; Loxodonta.
OX   NCBI_TaxID=9785;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=66120641; PubMed=5916966;
RA   Brown J.L., Hamilton J.B.;
RT   "Species variation in the amino acid sequence of insulin.";
RL   Am. J. Med. 40:662-666(1966).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: BOTH A-B (TWO CHAIN) AND A-A-B (THREE CHAIN) FORMS
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MISCELLANEOUS: THE SPECIES P. ELEPHAS IS NOT GIVEN, BUT IS
CC       MOST PROBABLY THE INDIAN ELEPHANT (ELEPHAS MAXIMUS).
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   PIR; A25941; INEL.
DR   HSSP; P01308; 1ZNI.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   PFINIS; PS00262; INSULINA.
DR   PSINIS; PS00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   ProSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_CONS      30    31
FT   CHAIN         31    51       INSULIN A CHAIN.
FT   DISULFID      7     47       INTERCHAIN.
FT   DISULFID      19    51       INTERCHAIN.
FT   DISULFID      31    47
SQ   SEQUENCE   51 AA;  5752 MW;  [illegible] CRC64;

  Query Match             93.0%;  Score 255.5;  DB 1;  Length 51;
  Best Local Similarity   94.2%;  Pred. No. 2.1e-24;
  Matches   49;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 50
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVEQCCASVCSLYQLENYCN 51


RESULT 3
INS_A[illegible]A
ID   INS_A[illegible]A      STANDARD;      PRT;    [illegible] AA.
AC   P[illegible];
DT   [illegible]-JUL-1988 (Rel. [illegible], Created)
DT   [illegible]-JUL-1988 (Rel. [illegible], Last sequence update)
DT   [illegible]-NOV-1997 (Rel. [illegible], Last annotation update)
```

```
DE   Insulin.
GN   INS.
OS   Acomys cahirinus (Egyptian spiny mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Acomys.
OX   NCBI_TaxID=10098;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=72029451; PubMed=5049113;
RA   Buse E.J., Dumbo, P.E.;
RT   "Isolation and partial structural analysis of insulin from [illegible]
RT   spinosissimus) and spiny mouse (Acomys cahirinus).";
RL   Hoppe-Seyler's Z. Physiol. Chem. 353:1441-1445(1972).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: BOTH A-B (TWO CHAIN) AND A-A-B (THREE CHAIN) FORMS
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   PIR; A25941; INEL.
DR   HSSP; P01308; 1ZNI.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   PFINIS; PS00262; INSULINA.
DR   PSINIS; PS00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   ProSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_CONS      30    31
FT   CHAIN         31    51       INSULIN A CHAIN.
FT   DISULFID      7     47       INTERCHAIN (BY SIMILARITY).
FT   DISULFID      19    51       INTERCHAIN (BY SIMILARITY).
FT   DISULFID      31    47       BY SIMILARITY.
SQ   SEQUENCE   51 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             91.5%;  Score 251.5;  DB 1;  Length 51;
  Best Local Similarity   [illegible]%;  Pred. No. 4.3e-24;
  Matches   48;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 50
           [illegible]
Db       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKARREVEDPQVEQLELGGGPGAGDLQTLALEVAQQKRGIVEQCCASVCSLYQLENYCN 51


RESULT 4
INS_CERAE
ID   INS_CERAE      STANDARD;      PRT;    110 AA.
AC   P30437; P01309;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   [illegible] FEB 1996 (Rel. 33, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Cercopithecus aethiops (Green monkey) (Grivet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Cercopithecus.
OX   NCBI_TaxID=9534;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92219953; PubMed=1560757;
RA   Gelfand C.A., Weiss S.J., Li W.;
RT   "Sequences of primate insulin genes support the hypothesis of a
RT   slower rate of molecular evolution in humans and apes than in
RT   monkeys.";
RL   Mol. Biol. Evol. 9:1-6(1992).
RN   [2]
RP   SEQUENCE OF [illegible].
RX   MEDLINE=[illegible]; PubMed=4[illegible];
RA   Patterson D.F., Nehrlich S., Iyer R.P., Steinet D.F.;
RT   "Determination of the amino acid sequence of the monkey, [illegible]
```

RX MEDLINE [illegible]; PubMed [illegible];
RA Schlor K., Vlik A.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 3. Synthesis of the sequences 14-17 and 9-13
RT of human proinsulin C peptides.";
RL Hoppe Seyler. Z. Physiol. Chem. 360:1205(1979).
RN [24]
RT SYNTHESIS OF RAT C.
RX MEDLINE [illegible]; PubMed [illegible];
RA Gattner K., Jaenner G., Kovatch W., Trauth G.;
RT "Synthesis of peptides with the properties of human proinsulin C
RT peptides (hC peptide). 1. Scheme for the synthesis and preparative
RT the sequence 28-31 of human proinsulin C peptide.";
RL Hoppe Seyler. Z. Physiol. Chem. 358:1183(1977).
RN [25]
RP VARIANT LOS ANGELES.
RX MEDLINE [illegible]; PubMed [illegible];
RA Haneda M., Chan S.J., Kwok S.C.M., Rubenstein A.H., Steiner D.F.;
RT "Studies on mutant human insulin genes: identification and sequence
RT analysis of a gene encoding [SerB24]insulin.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:6366-6370(1983).
RN [26]
RP VARIANTS LOS ANGELES AND CHICAGO.
RX MEDLINE [illegible]; PubMed [illegible];
RA Shoelson S., Pickova M., Haneda M., Nahum A., Musso G., Kaiser E.T.,
RA Rubenstein A.H., Tager H.;
RT "Identification of a mutant human insulin predicted to contain a
RT serine-to-phosphotyrosine substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3798-3801(1983).
RN [27]
RP VARIANT CHICAGO.
RX MEDLINE [illegible]; PubMed [illegible];
RA Shoelson S., Haneda M., Blix P., Nanjo K., Sanke T., Inouye K.,
RA Steiner D.F., Rubenstein A.H., Tager H.;
RT "A mutation in the B chain coding region is associated with impaired
RT proinsulin conversion in a family with hyperproinsulinemia.";
RL J. Clin. Invest. 72:732-747(1983).
RN [28]
RP VARIANT WAKAYAMA.
RX MEDLINE [illegible]; PubMed [illegible];
RA Sakura H., Inoue T., Yamamoto Y., Kobayashi T., Hirata H.,
RT "Structurally abnormal insulin in a diabetic patient. Characterization
RT of the mutant insulin Asn (Val) isolated from the pancreas.";
RL J. Clin. Invest. 78:1660-1672(1986).
RN [29]
RP VARIANT HIS65.
RX MEDLINE [illegible]; PubMed [illegible];
RA Barbetti F., Raben N., Kadowaki T., Cama A., Accili D., Taylor K.R.,
RA Moses A.M., Joyce S.C., Roth J.;
RT "Two unrelated patients with familial hyperproinsulinemia due to a
RT mutation substituting histidine for arginine at position 65 in the
RT proinsulin molecule: identification of the mutation by direct
RT sequencing of genomic deoxyribonucleic acid amplified by polymerase
RT chain reaction.";
RL J. Clin. Endocrinol. Metab. 71:1401-1406(1990).
RN [30]
RP VARIANT HIS65.
RX MEDLINE [illegible]; PubMed [illegible];
RA Shibasaki Y., Kawakami T., Kanazawa Y., Ishida Y., Tamura T.,
RT "Posttranslational cleavage of proinsulin is blocked by a point
RT mutation in familial hyperproinsulinemia.";
RL J. Clin. Invest. 76:578-583(1985).
RN [31]
RP VARIANT LYS [illegible].
RX MEDLINE [illegible]; PubMed [illegible];
RA Yano H., Kitano N., Morimoto M., Polonsky K.S., Imura H., Seino Y.;
RT "A novel point mutation in the human insulin gene giving rise to
RT hyperproinsulinemia (proinsulin Kyoto).";
RL J. Clin. Invest. Medicine 13:67-72(1990).
RN [32]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q.X., Weiss M.A.;
RT "Toward the solution structure of human insulin: sequential 2D 1H NMR
RT assignment of a des-pentapeptide mini-B and comparison with native
RT structure.";
RL Biochemistry [illegible] [illegible].
RN [33]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q.X., Weiss M.A.;
RT "Comparative 2D NMR studies of human insulin and des-pentapeptide
RT insulin: sequential resonance assignment and implications for protein
RT dynamics and receptor recognition.";
RL Biochemistry [illegible] [illegible].
RN [34]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q.X., Weiss M.A.;
RT "Two-dimensional NMR studies of bis(B26 B29) insulins: sequence
RT specific resonance assignments and effects of solvent exposure.";
RL Biochim. Biophys. Acta [illegible] [illegible].
RN [35]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q.X., Weiss M.A.;
RT "Three-dimensional structure of an insulin dimer. A study of
RT the B9(Asp) mutant of human insulin using nuclear magnetic resonance,
RT distance geometry and restrained molecular dynamics.";
RL J. Mol. Biol. 227:1143-1163(1992).
RN [36]
RP STRUCTURE BY NMR OF VARIANT LYS ANGELES.
RX MEDLINE [illegible]; PubMed [illegible];
RA Hua Q.X., Shoelson S.E., Inouye K., Weiss M.A.;
RT "Paradoxical structure and function in a mutant human insulin
RT associated with diabetes mellitus.";
RL Proc. Natl. Acad. Sci. U.S.A. [illegible] [illegible].
RN [37]
RP STRUCTURE BY NMR.
RX MEDLINE [illegible]; PubMed [illegible];
RA Chua X., Bentley G., Dodson E.J., Dodson G.G.;

Query Match [illegible]%; Score 267; DB 2; Length [illegible];
Best Local Similarity [illegible]%; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 2
INS_MACFA
ID INS_MACFA STANDARD; PRT; [illegible] AA.
AC P30406; P01309;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-AUG-1997 (Rel. 35, Last sequence update)
DT 01-FEB-1996 (Rel. 34, Last annotation update)
DE Insulin precursor.
GN INS.
OS Macaca fascicularis (Crab-eating Macaque) (Cynomolgus Monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
OX NCBI Taxid=9541;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE [illegible]; PubMed [illegible];
RA Wetekam W., Steinbeck J., Lehmacher M., Wendenmayer F.,
RA Winnacker E.-L.;
RT "The nucleotide sequence of cDNA coding for proinsulin from the
RT primate Macaca fascicularis.";
RL Gene 17:79-84(1982).

```
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   PIR; A01582; INWHIS.
DR   HSSP; P01317; 1APH.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_CONS     30     31
FT   CHAIN        31     51       INSULIN A CHAIN.
FT   DISULFID      7     37       INTERCHAIN.
FT   DISULFID     19     50       INTERCHAIN.
FT   DISULFID     36     41
SQ   SEQUENCE   51 AA;  572[illegible] MW;  9807B[illegible] CRC64;

  Query Match             89.6%;  Score 263.5;  DB 1;  Length 51;
  Best Local Similarity   92.3%;  Pred. No. [illegible];
  Matches   48;  Conservative    0;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
Db          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKA-GIVEQCCAS[illegible]LYQLENYCN 51


RESULT 9
INS_CAMDR
ID   INS_CAMDR      STANDARD;      PRT;    51 AA.
AC   P01320;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Insulin.
GN   INS.
OS   Camelus dromedarius (Dromedary) (Arabian camel).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX   NCBI_TaxID=9838;
RN   [1]
RP   SEQUENCE.
RA   Danho W.O.;
RT   "The isolation and characterization of insulin of camel (Camelus
RT   dromedarius).";
RL   J. Fac. Med. Baghdad 14:1[illegible]-28(1972).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   PIR; A92782; INCMA.
DR   HSSP; P01317; 2INS.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_CONS     30     31
FT   CHAIN        31     51       INSULIN A CHAIN.
FT   DISULFID      7     37       INTERCHAIN.
FT   DISULFID     19     50       INTERCHAIN.
FT   DISULFID     36     41
SQ   SEQUENCE   51 AA;  569[illegible] MW;  [illegible] CRC64;

  Query Match             89.6%;  Score 263.5;  DB 1;  Length 51;
  Best Local Similarity   90.4%;  Pred. No. 7.3e-25;
  Matches   47;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
Db          1 FANQHLCGSHLVEALYLVCGERGFFYTPKA-GIVEQCCAGVCSLYQLENYCN 51


RESULT 10
INS_CAPHI
ID   INS_CAPHI      STANDARD;      PRT;    51 AA.
AC   P01319;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   Insulin.
GN   INS.
OS   Capra hircus (Goat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; [illegible]
OC   Bovidae; Caprinae; Capra.
OX   NCBI_TaxID=9925;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE 66168119; PubMed 5949[illegible];
RA   Smith L.F.;
RT   "Species variation in the amino acid sequence of insulin.";
RL   Am. J. Med. 40:662-666(1966).
CC   -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR   PIR; A01586; INGT.
DR   HSSP; P01317; 1APH.
DR   InterPro; IPR004825; Ins/IGF/relax.
DR   PRINTS; PR00276; INSULINA.
DR   PRINTS; PR00277; INSULINB.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism.
FT   CHAIN         1     30       INSULIN B CHAIN.
FT   NON_CONS     30     31
FT   CHAIN        31     51       INSULIN A CHAIN.
FT   DISULFID      7     37       INTERCHAIN.
FT   DISULFID     19     50       INTERCHAIN.
FT   DISULFID     36     41
SQ   SEQUENCE   51 AA;  5662 MW;  9007B[illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB 1;  Length 51;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   47;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRGIVEQCCTSICSLYQLENYCN 52
Db          1 [illegible]LENYCN 51


RESULT 11
INS_PIG
ID   INS_PIG        STANDARD;      PRT;   108 AA.
AC   P01315; [illegible];
DT   21-JUL-1986 (Rel. 01, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RA   Ban X.-L., Lich B.E.;
RT   "Complete porcine preproinsulin cDNA sequence.";
RL   Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 25-109.
RX   MEDLINE=82286187; PubMed=7172706;
RA   Chance R.E., Ellis R.M., Bromer W.W.;
RT   "Porcine proinsulin: characterization and amino acid sequence.";
RL   Science 161:165-167(1968).
RN   [3]
RP   REVISION OF 59.
RA   Chance R.E.;
RL   Submitted (JUL-1970) to the PIR data bank.
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA   Brandenburg D., Gattner H.-G., Jorns A., Wollmer A.;
RT   "Monomeric insulins: characterization and molecular properties.";
RL   Horm. Metab. Res. Suppl. 24:11-18(1989).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RA   Brzovic P., Dodson E., Dodson G., Meng W., Hu S.-H., Mackin R.B.,
RA   Steiner D.F.;
RT   "Structural studies of human insulin: 1.9 A crystal structure of
RT   human insulin and 2.0 A structure of human des(B26-B30) insulin.";
RL   J. Mol. Biol. 239:774-790(1994).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RA   Brzovic P., Dodson E., Dodson G., Meng W., Hu S.-H., Mackin R.B.,
RA   Steiner D.F.;
RT   "Structural studies of human insulin: 1.9 A crystal structure of
RT   human insulin and 2.0 A structure of human des(B26-B30) insulin.";
RL   J. Mol. Biol. 239:774-790(1994).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=94120280; PubMed=7506311;
RA   Ciszak E., Smith G.D.;
RT   "Crystallographic evidence of insulin flexibility in an acidic
RT   environment.";
RL   Biochemistry 33:1512-1517(1994).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).
RX   MEDLINE=97099418; PubMed=8989485;
RA   Baker E.N., Blundell T.L., Cutfield J.F., Cutfield S.M., Dodson E.J.,
RA   Dodson G.G., Crowfoot Hodgkin D.M., Hubbard R.E., Isaacs N.W.,
RA   Reynolds C.D., Sakabe K., Sakabe N., Vijayan N.M.;
RT   "The structure of 2Zn pig insulin crystals at 1.5 A resolution.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 319:369-456(1988).
RN   [9]
RP   X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX   MEDLINE=94120280; PubMed=7506311;
RA   Ciszak E., Smith G.D.;
RT   "Crystallographic evidence of insulin flexibility in an acidic
RT   environment.";
RL   Biochemistry 33:1512-1517(1994).
RN   [10]
RP   X-RAY CRYSTALLOGRAPHY.
RX   MEDLINE=91222450; PubMed=2025415;
RA   Badger J., Harris M.R., Reynolds C.D., Evans A.C., Dodson E.J.,
RA   Dodson G.G., North A.C.T.;
RT   "Structure of the pig insulin dimer in the cubic crystal.";
RL   Acta Crystallogr. B 47:127-136(1991).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (1.65 ANGSTROMS).
RA   Hua Q.-X., Wan Z.-L., Chance M.R., Liang D.-C.;
RT   "Structure of monomeric porcine DesB1 B2 desoctapeptide (B25-B32)
RT   insulin at 1.65-A resolution.";
RL   Acta Crystallogr. D 53:507-512(1997).
CC   -!- FUNCTION: INSULIN INCREASES BLOOD GLUCOSE CONCENTRATION. IT
CC       INCREASES CELL PERMEABILITY TO MANY AMINO ACIDS, AND
CC       FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC       CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC   -!- SUBUNIT: HEXAMER. A DIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC       DISULFIDE BONDS.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC   -!- DATABASE: NAME=Protein Spotlight;
CC       NOTE=Issue 9 of April 2001;
CC       WWW="http://www.expasy.ch/spotlight/spotlight-9.html"
CC   ---------- ---------- ---------- ---------- ----------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
DR   EMBL; AF004555; AAC27926.1; ALT_INIT.
DR   PIR; A01584; IPIG.
DR   PDB; 1BEN; 06-JUL-94.
DR   PDB; 4INS; 31-JUL-94.
DR   PDB; 6INS; 31-JAN-94.
DR   PDB; 7INS; 31-JAN-94.
DR   PDB; 9INS; 15-APR-94.
DR   PDB; 1IZA; 15-APR-94.
DR   PDB; 1IZB; 15-APR-94.
DR   PDB; 2PIC; 29-JAN-98.
DR   PDB; 1MPJ; 29-JAN-98.
DR   PDB; 3INS; 29-JAN-98.
DR   PDB; 1EEI; 15-JUN-94.
DR   PDB; 1SDB; 31-AUG-95.
DR   PDB; 1WAV; 28-FEB-97.
DR   PDB; 1ZEI; 31-MAR-96.
DR   PDB; 1ZNI; 28-JAN-98.
DR   PDB; 1ZNJ; 28-JAN-98.
DR   InterPro; IPR003825; Ins/IGF/relax.
DR   Pfam; PF00049; Insulin; 1.
DR   PRINTS; PR00276; INSULINGNA.
DR   PRINTS; PR00277; INSULINE.
DR   SMART; SM00078; IlGF; 1.
DR   PROSITE; PS00262; INSULIN; 1.
KW   Insulin family; Hormone; Glucose metabolism; Signal; Disulfide bond.
FT   SIGNAL       1     24
FT   CHAIN       25     54       INSULIN B CHAIN.
FT   PROPEP      57     87       C PEPTIDE.
FT   CHAIN       88    109       INSULIN A CHAIN.
FT   DISULFID     31     91       INTERCHAIN.
FT   DISULFID     43    107       INTRACHAIN.
FT   DISULFID     91    107
FT   HELIX        29     36
FT   STRAND       40     45
FT   HELIX        49     54
FT   HELIX        90     94
FT   STRAND      100    105
FT   STRAND      102    107
SQ   SEQUENCE   109 AA;  11937 MW;  3A6F1C812C0CBE0C CRC64;

  Query Match             85.5%;  Score 2553;  DB 1;  Length 109;
  Best Local Similarity   94.7%;  Pred. No. 7.4e-276;
  Matches   51;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT 31
             ||||||||||||||||||||||||||||||
Db      25 FVNQHLCGSHLVEALYLVCGERGFFYTPKT 55

Qy      31 RREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 85
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      85 QKRGIVEQCCTSICSLYQLENYCN 109


RESULT 12
ING_HAB11
ID   ING_HAB11               STANDARD;      PRT;   110 AA.
AC   P01311;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   Insulin precursor.
GN   INS.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Lagomorpha; Leporidae;
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=New Zealand white; TISSUE=Pancreas;
RX   MEDLINE=84179230; PubMed=6143757;
RA   Deshpande S.S., Gibbons S.J., Rajbhandary U.A., Khorana H.G.;
RA   Moore W.R., Cahn J.S.;
RT   "Insulin gene expression and insulin synthesis in mammalian cells and
```

Title: US-09-423-100-5
Perfect score: 264
Sequence: 1 EVNLRLYGSHLVEALYLVCG.....IVEQCCTSICSLYQLENYCN 52

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 621580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 621580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_synthetic:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaps:*
17: sp_archeaps:*

Prob. N.. is the number of results probability chance. Expect is the number of hits you can expect to obtain in this database search with a score at least as good as the result being printed. Total is followed by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	264	100.0	52	21	_qIXIs	_qIXIs spermophilu
2	262	99.4	52	4	_qWNW7	qWNW7 Ictidomys t
3	262.5	99.4	52	21	_qIZB_7	_qIZB_7 Ictidomys
4	261.5	99.2	52	21	_qIAM	_qIAM Ictidomys
5	261.5	99.2	52	21	_qIEE	_qIEE Ictidomys
6	261.5	99.2	52	21	_qI ZN4	qI ZN4 Callithri
7	196	74.3	111	21	_qIBB	qIBB Chlorocebus
8	194.5	73.8	110	21	_qIOZY	qIOZY Rhodon alos
9	192.5	73.0	111	21	_qIAT	qIAT Osteoglossu
10	189.5	71.8	87	21	_qIAA	_qIAA Anabaenoms
11	189.5	71.8	104	21	_qIAC	_qIAC Nostoc sp.
12	189.5	71.8	91	21	_qIB2	_qIB2 Arthrospir
13	188	71.2	47	21	_qIXD	_qIXD Spirulina
14	183	69.3	225	21	_qIZ2	_qIZ2 Cyanothece
15	182.5	69.1	95	21	_qIC7	_qIC7 Parachlorb
16	182.5	69.1	102	21	_qIZ24	_qIZ24 Oreochromis
17	181.5	68.8	99	21	_qIC_	_qIC_ Cyanothece
18	181.5	68.8	99	21	_qIC0	_qIC0 Cyanothece
19	181.5	68.8	99	21	_qIC1	_qIC1 Cyanothece
20	181.5	68.8	99	21	_qIC2	_qIC2 Cyanothece
21	181.5	68.8	99	21	_qIC3	_qIC3 Cyanothece
22	181.5	68.8	99	21	_qIC4	_qIC4 Cyanothece
23	181.5	68.8	99	21	_qIC5	_qIC5 Cyanothece
24	181.5	68.8	99	21	_qIC6	_qIC6 Cyanothece
25	181.5	68.8	99	21	_qIC8	_qIC8 Cyanothece
26	181.5	68.8	99	21	_qIC9	_qIC9 Cyanothece
27	181.5	68.8	99	21	_qICA	_qICA Cyanothece
28	181.5	68.8	99	21	_qICB	_qICB Cyanothece
29	181.5	68.8	99	21	_qICC	_qICC Cyanothece
30	181.5	68.8	99	21	_qICD	_qICD Cyanothece
31	181.5	68.8	99	21	_qICE	_qICE Cyanothece
32	181.5	68.8	99	21	_qICF	_qICF Cyanothece
33	181.5	68.8	99	21	_qICG	_qICG Cyanothece
34	181.5	68.8	99	21	_qICH	_qICH Cyanothece
35	181.5	68.8	99	21	_qICI	_qICI Cyanothece
36	181.5	68.8	99	21	_qICJ	_qICJ Cyanothece
37	181.5	68.8	99	21	_qICK	_qICK Cyanothece
38	181.5	68.8	99	21	_qICL	_qICL Cyanothece
39	181.5	68.8	99	21	_qICM	_qICM Cyanothece
40	181.5	68.8	99	21	_qICN	_qICN Cyanothece
41	181.5	68.8	99	21	_qICO	_qICO Cyanothece
42	181.5	68.8	99	21	_qICP	_qICP Cyanothece
43	181.5	68.8	99	21	_qICQ	_qICQ Cyanothece
44	181.5	68.8	99	21	_qICR	_qICR Cyanothece
45	181.5	68.8	99	21	_qICS	_qICS Cyanothece

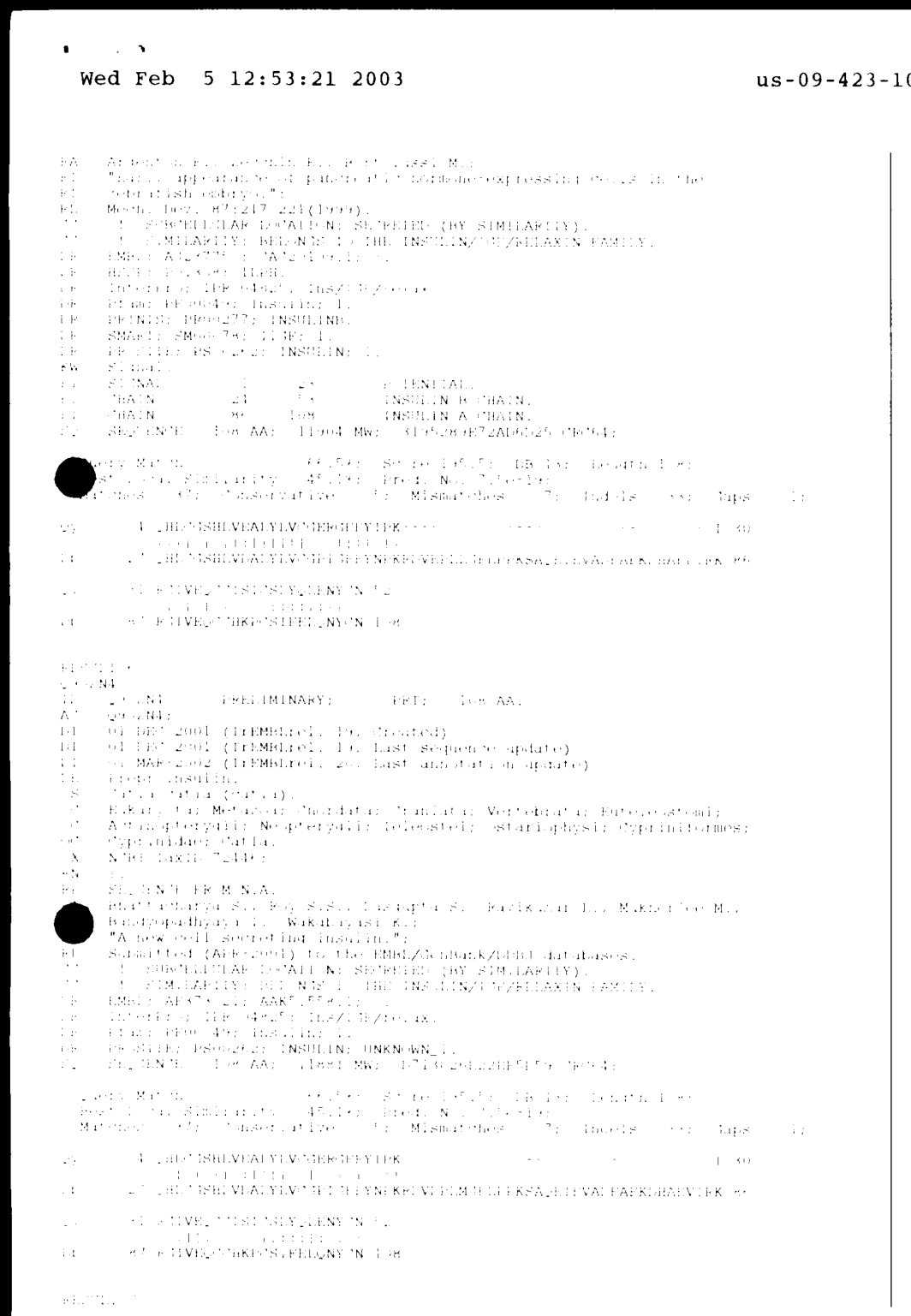
ALIGNMENTS

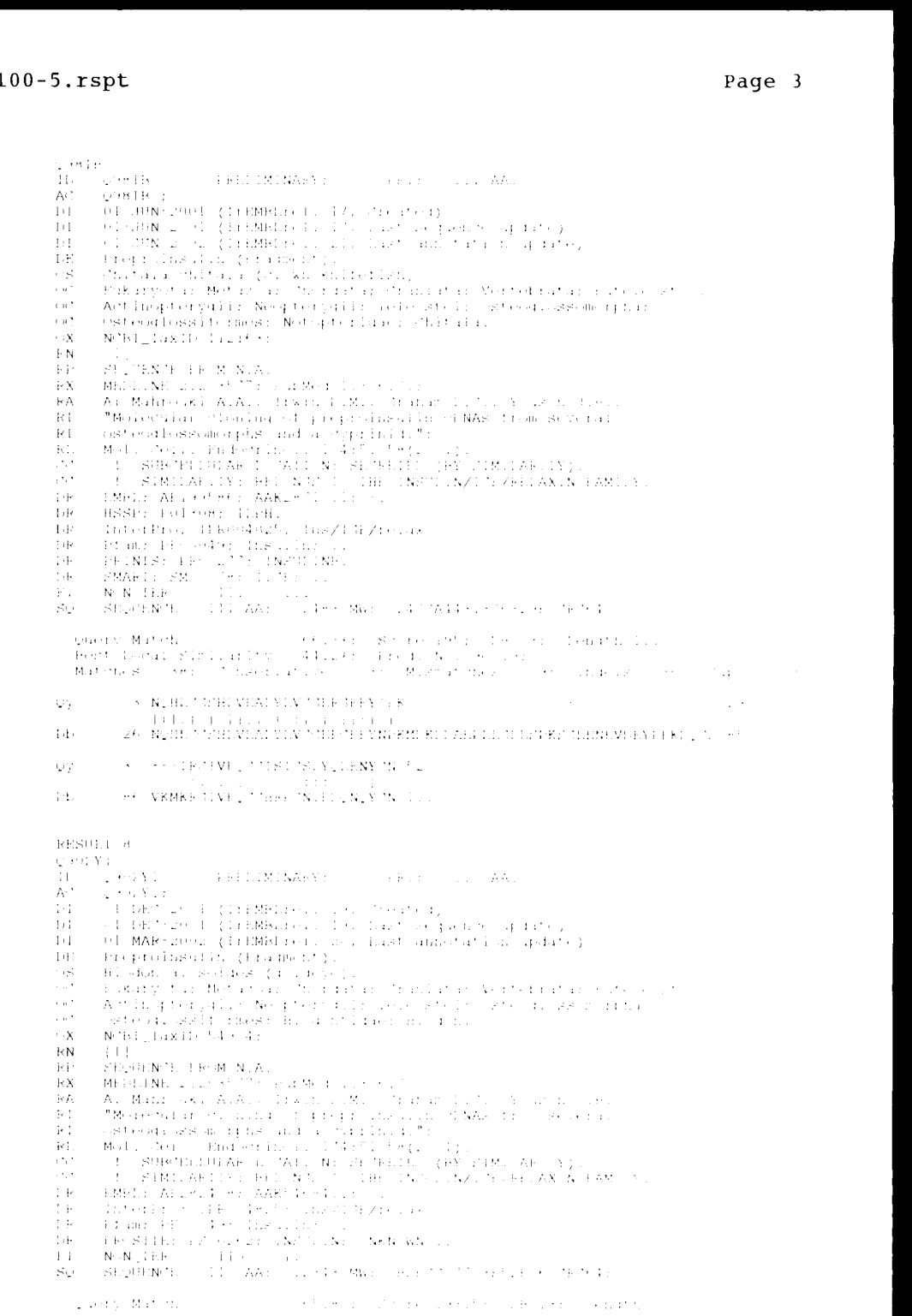
RESULT 1
qIXIs
ID qIXIs PRELIMINARY; PRT; 52 AA.
AC qIXIs;
DT 01-FEB-2003 (TrEMBLrel. 21, Created)
DT 01-FEB-2003 (TrEMBLrel. 21, Last sequence update)
DT 15-MAR-2003 (TrEMBLrel. 22, Last annotation update)
DE Insulin.
GN Spermophilus tridecemlineatus (thirteen-lined ground squirrel).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Spermophilus.
OX NCBI_TaxID=43179;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREAS;
RA Lesniak M.M., Rios M.S., Tamarit E.S., Stock I.I., Agricola M.;
RL "Beta-cell loss and insulitis in hibernating ground squirrels.";
RL Submitted (JUN-2002) to EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE INSULIN FAMILY.
DR EMBL; AY080604; AAK72[illegible].
DR InterPro; IPR004825; Insulin.
DR Pfam; PF00049; Insulin.
DR PROSITE; PS00262; INSULIN.
SQ SEQUENCE 52 AA; [illegible] MW; [illegible] CRC32;

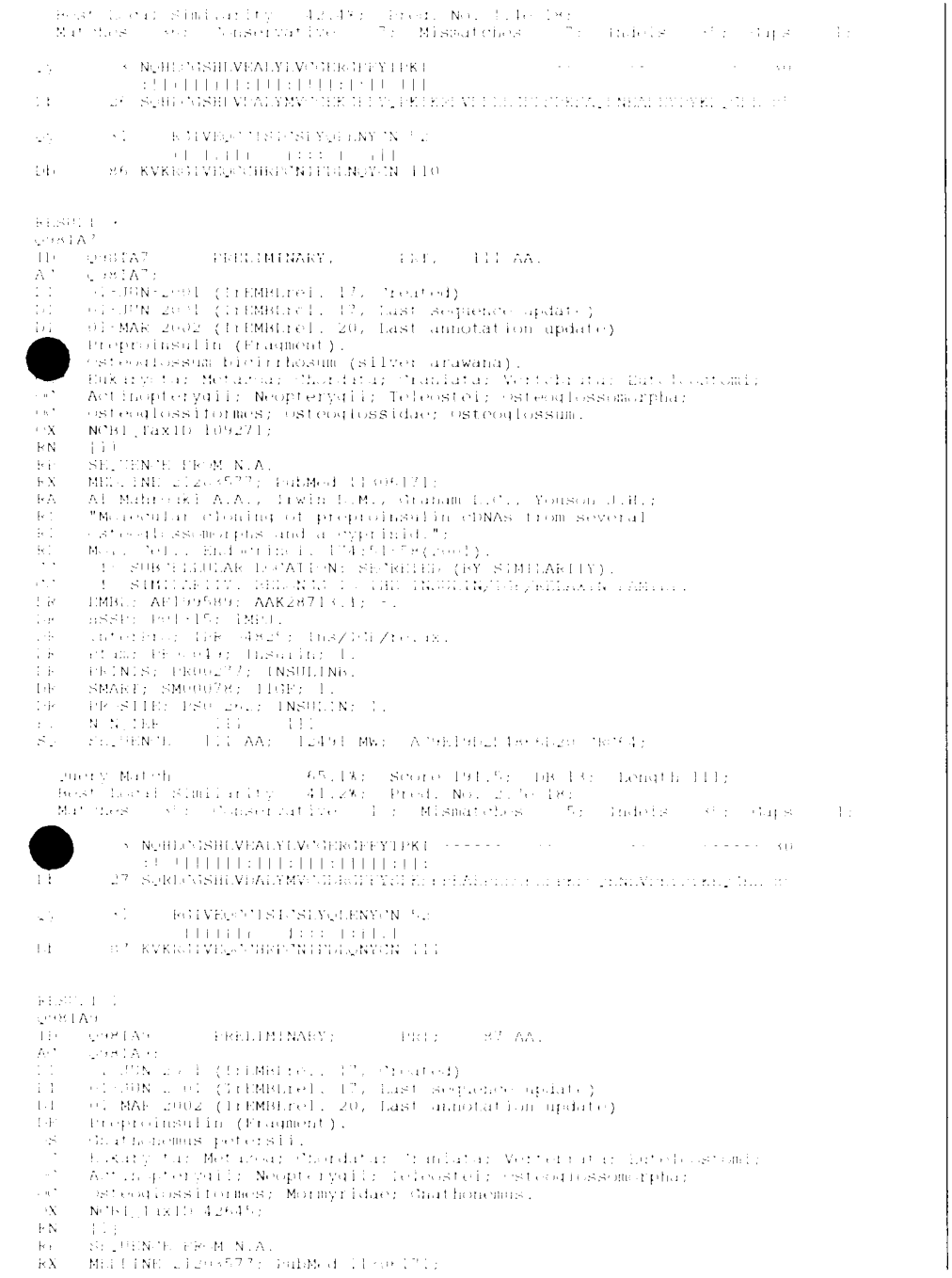
Query Match [illegible] Score [illegible]; DB [illegible]; Length 52;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy 1 EVNLRLYGSHLVEALYLVCG 20
[illegible]
Db [illegible] EVNLRLYGSHLVEALYLVCG [illegible]

Qy [illegible]
[illegible]
Db [illegible] AlaEssIVEQCCTSICSLYQLENYCN [illegible]







RESULT 12
Q95182
ID Q95182 PRELIMINARY; PRT; 91 AA.
AC Q95182;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproinsulin (Fragment).
OS Amblyopites rupestris.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidei;
OC Centrarchidae; Amblopilites.
OX NCBI_TaxID=39273;
RN [1]
RP SEQUENCE FROM N.A.
RA Al-Maleki A.A., Irwin D.M., Yeomans J.B.;
RT "Molecular cloning of preproinsulin cDNA from the rock bass.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/INSULIN-LIKE GROWTH
DR EMBL; AF199594; AAK28708.1; -.
DR HSSP; P01308; 1LPH.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
FT NON_TER 1 1
FT NON_TER 91 91
SQ SEQUENCE 91 AA; 10100 MW; E9608B2563B2691B9 CRC64;

Query Match 64.1%; Score 185.5; DB 17; Length 91;
Best Local Similarity 40.9%; Pred. No. 2.4e-17;
Matches 36; Conservative 5; Mismatches 8; Indels 9; Gaps 1;

Qy 4 LVELGGSHLVEALYLVCGERGFFYTPK 29
||||||||||||||||||||||| ||
Db 4 LVELGGSHLVEALYLVCGERGFFYTPK 27

Qy 30 ----IEGIVEQCCTSICSLYQLENYCN 52
|||||||| |||| | |||
Db 28 RLMMVKEGIVEQCCHHPCNIFELQKYCN 91

RESULT 13
Q9PX06
ID Q9PX06 PRELIMINARY; PRT; 207 AA.
AC Q9PX06;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproinsulin-like growth factor-II.
GN IGF-II.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee M.C.L., Chan K.M., Chen C.H.K.;
RT "Fish cloning and Gene Expression Studies on Common Carp (Cyprinus
RT carpio) Insulin-like Growth Factor-II.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AF302958; AAL25799.1; -.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PROSITE; PS00262; INSULIN; UNKNOWN_1.

RESULT 14
Q73721
ID Q73721 PRELIMINARY; PRT; 27 AA.
AC Q73721;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Insulin-like growth factor II.
GN IGF-II.
OS Tilapia sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC Cichlidae; Tilapia.
OX NCBI_TaxID=8279;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen S.Y., Chang C.Y., Chou C.C., Shen S.C., Wu J.L.;
RT "Production of biologically active recombinant tilapia insulin-like
RT growth factor II polypeptides in E. coli cells and characterization
RT the genomic structure of the coding region.";
RL DNA Cell Biol. 0:0-0(1998).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AF079821; AAC79495.1; -.
DR HSSP; P01344; 1GF2.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 215 AA; 24719 MW; [illegible] CRC64;

Query Match 34.1%; Score 98.5; DB 17; Length 215;
Best Local Similarity 40.2%; Pred. No. 1.4e-11;
Matches 37; Conservative 7; Mismatches 12; Indels 9; Gaps 1;

Qy 6 LCGSHLVEALYLVCGERGFFYTPKT 33
||| |||| |||||| |||| |||||||
Db 55 LCGGELVEALQFVCGDRGFYFSRPAS 81

RESULT 15
Q9W6Q7
ID Q9W6Q7 PRELIMINARY; PRT; 179 AA.
AC Q9W6Q7;
DT 01-NOV-1999 (TrEMBLrel. 8, Created)
DT 01-NOV-1999 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Preproinsulin-like growth factor Ia.
GN IGF-I.
OS Paralichthys olivaceus (Bastard).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectoidei;
OC Pleuronectoidei; Paralichthyidae; Paralichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim D.S.;
RT "Expression of IGF-I cDNA fused to [illegible]

RT in mammalian CHO cell line using green fluorescence protein (GFP)
RT tagging: secretory production of big IGF1b-GFP fusion proteins from
RT stable transfected CHO cell culture.";
RL Submitted (XXX-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
DR EMBL; AJ010602; CAA09267.1; -.
DR HSSP; P01343; 1GF1.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00276; INSULINA.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; IlGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
SQ SEQUENCE 159 AA; 17541 MW; 8B611C38583E0865 CRC64;

Query Match 48.1%; Score 141.5; DB 13; Length 159;
Best Local Similarity 47.3%; Pred. No. 2.2e-11;
Matches 26; Conservative 8; Mismatches 12; Indels 9; Gaps 1;

Qy 5 LVGSHLVEALYLVCGERGFFYTPKT---------KRDVEDPQVEQLELGGSLYQLENYC 51
Db 19 LVGAELVDALQFVCGDRGFYFNKPT---------GYGSSSRRAPQTGIVDECCFRSCDLRRLEMYC 73

Search completed: February 5, 2003, 12:47:30
Job time : 17.9606 secs

Title: US-09-423-100-5

Perfect score: 294

Sequence: 1 MVSLHEDSHLVEAIYIVI.....IVEQVISQSIV[1]MVSN 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2647 seqs, 2042522 residues

Total number of hits satisfying chosen parameters: 2647

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/ina/PCTUS_*.pep:*
2: /cgn2_6/ptodata/2/ina/PCTUS_*.pep:*
3: /cgn2_6/ptodata/2/ina/PCTUS_*.pep:*
4: /cgn2_6/ptodata/2/ina/PCTUS_*.pep:*
5: /cgn2_6/ptodata/2/ina/backfile1.pep:*
6: /cgn2_6/ptodata/2/ina/backfile2.pep:*

Query Match: Score - Result No.; Length of Query
Best Local Similarity: Percent Similarity
Matches: Number of Matches; Conservative: Number of Mismatches; Indels: Number of Gaps

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	294	100.0	96	1	US-08-160-676A-7	Sequence 7, Appli
2	294	100.0	96	1	US-08-***-***A-11	Sequence 11, Appli
3	294	100.0	96	1	US-08-***-***A-2	Sequence 2, Appli
4	294	100.0	96	1	US-08-291-676A-5	Sequence 5, Appli
5	294	100.0	96	1	US-08-160-676A-5	Sequence 5, Appli
6	294	100.0	96	1	US-08-***-***-8	Sequence 8, Appli
7	294	100.0	147	1	US-08-***-***-59	Sequence 59, Appl
8	294	100.0	147	1	US-08-***-***-59	Sequence 59, Appl
9	294	100.0	148	1	US-08-***-***-45	Sequence 45, Appl
10	294	100.0	148	1	US-08-***-***-45	Sequence 45, Appl
11	294	100.0	148	1	US-08-406-256-48	Sequence 48, Appl
12	294	100.0	148	1	US-08-***-***-48	Sequence 48, Appl
13	294	100.0	57	1	US-***-***-7-1A-44	Sequence 44, Appl
14	294	100.0	96	1	US-***-***-17-4	Sequence 4, Appli
15	294	100.0	96	4	US-***-***-***A-42	Sequence 42, Appl
16	294	100.0	97	4	US-***-***-***A-1	Sequence 1, Appli
17	294	100.0	96	1	US-08-258-617-3	Sequence 3, Appli
18	294	100.0	96	4	US-08-***-***-3	Sequence 3, Appli
19	294	100.0	96	4	US-08-***-***-4	Sequence 4, Appli
20	294	100.0	96	4	US-08-***-***-5	Sequence 5, Appli
21	294	100.0	96	4	US-08-***-***A-1	Sequence 1, Appli
22	294	100.0	96	4	US-08-***-***A-5	Sequence 5, Appli
23	294	100.0	97	4	US-08-***-***-7	Sequence 7, Appli
24	275.5	93.7	58	4	US-09-261-***-2	Sequence 2, Appli
25	275.5	93.7	65	1	US-08-468-674B-71	Sequence 71, Appl
26	275.5	93.7	65	1	US-08-780-671-71	Sequence 71, Appl
27	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
28	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
29	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
30	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
31	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
32	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
33	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
34	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
35	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
36	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
37	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
38	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
39	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
40	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
41	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
42	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
43	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
44	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl
45	275.5	93.7	84	1	US-08-***-***-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-160-676A-7
; Sequence 7, Application US/08160676A
; Patent No. 5773571
; GENERAL INFORMATION:
; APPLICANT: Holmes, Kathryn V.
; APPLICANT: Dveksler, Gabriel
; APPLICANT: Cardellichio, Vincent E.
; APPLICANT: Gorny, Maria
; TITLE OF INVENTION: Receptor for Murine Coronavirus
; TITLE OF INVENTION: Soluble Form Thereof
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenneth A. Gallo, Esq.
; STREET: 1901 L Street, N.W., Suite 500
; CITY: Washington
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 20036-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,676A
; FILING DATE: December 1, 1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/468,674
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara V. Mercer, Esq.
; REGISTRATION NUMBER: 31,287
; REFERENCE/DOCKET NUMBER: 0112.0003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 331-7111
; TELEFAX: (202) 293-6229
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 96 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Unknown
US-08-160-676A-7

Query Match 100.0%; Score 294; DB 1; Length 96;
Best Local Similarity 100.0%; Pred. No. 2.4e-29;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10171-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100-2
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 7067.122-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-423-100-49

Query Match 100.0%; Score 294; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 137

RESULT 9
US-08-400-256-49
; Sequence 49, Application US/08400256
; Patent No. 5714447
; GENERAL INFORMATION:
; APPLICANT: Bayelund, Svend
; APPLICANT: Balstrom, John
; APPLICANT: Jonassen, IB
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 7067.122-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-400-256-49

Query Match 100.0%; Score 294; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 137

; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-423-100-49

Query Match 100.0%; Score 294; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 137

RESULT 9
US-08-400-256-49
; Sequence 49, Application US/08400256
; Patent No. 5714447
; GENERAL INFORMATION:
; APPLICANT: Bayelund, Svend
; APPLICANT: Balstrom, John
; APPLICANT: Jonassen, IB
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 7067.122-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 137 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-400-256-49

Query Match 100.0%; Score 294; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 1.3e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREVEDPQVGQLELGGGPGAGSLQPLALEGSLQKRGIVEQCCTSICSLYQLENYCN 137

RESULT 10
US-08-975-865-49

; Sequence 45, Application US/08975365
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible] MAR 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-975-365-45

Query Match [illegible]; Score 294; DB [illegible]; Length 145;
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] FVNQHLCGSHLVEALYLV[illegible] 52

Db 94 FVNQHLCGSHLVEALYLVCGERGFFYTPKT[illegible] 145

RESULT [illegible]
[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/[illegible]
; FILING DATE: [illegible] MAR 1997
; CLASSIFICATION: [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible] amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-[illegible]

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible] FVNQHLCGSHLVEALYLV[illegible]

Db [illegible] FVNQHLC[illegible]

RESULT 12
US-08-975-365-48
; Sequence 48, Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: [illegible]
; STREET: 405 Lexington Avenue, [illegible]
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: [illegible]
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/[illegible]
; FILING DATE:
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible] MAR [illegible]
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: [illegible]
; TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
; SEQUENCE CHARACTERISTICS:
; LENGTH: [illegible] amino acids

; Type: amino acid
; Topology: linear
; Molecule Type: protein
; [illegible]

Query Match 100.0%; Score 294; DB 5; Length 146;
Best Local Similarity 100.0%; Prod. No. 1.1e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLVSSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 52
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 94 EVNQHLVSSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVARQKRGIVDQCCTSICSLYQLENYCN 145

RESULT 14
; Sequence 44, Application US/08630741A
; Patent No. 5426036
; GENERAL INFORMATION:
; APPLICANT: Keller, Klaus Peter
; APPLICANT: Kloss, Günther Johannes
; APPLICANT: Ohlmann, Edgar
; APPLICANT: Wallmeier, Helger
; TITLE OF INVENTION: Processes for the Preparation of Foreign
; TITLE OF INVENTION: Proteins in Streptomycetes
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/[illegible]
; FILING DATE: 12-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/[illegible]
; FILING DATE: 08 MAY 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/430,022
; FILING DATE: 01-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/[illegible]
; FILING DATE: 19 APR 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/735,777
; FILING DATE: 29-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07 11 866,4
; FILING DATE: 05-MAY-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 38 37 27[illegible]
; FILING DATE: 08-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 39 27 449.7
; FILING DATE: 19-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/[illegible]
; FILING DATE: 21-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Klauber, Michael K.
; REGISTRATION NUMBER: 34,851
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 408 4000
; TELEFAX: 202 408 4400
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-630-741A-44

Query Match [illegible]; Score [illegible]; DB 5; Length [illegible];
Best Local Similarity [illegible]; Prod. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
[illegible]
Db [illegible]

RESULT 14
US-08-233-617-4
; Sequence 4, Application US/08[illegible]
; Patent No. 5466666
; GENERAL INFORMATION:
; APPLICANT: Obermeier, Rainer
; APPLICANT: Sabel, Walter
; APPLICANT: Seitz, Peter
; APPLICANT: Geisert, Karl
; TITLE OF INVENTION: Amorphous Monospherical Forms of Basic
; TITLE OF INVENTION: Derivatives
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/233,617
; FILING DATE: 25-APR-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 43 13 [illegible]
; FILING DATE: 27-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Carol P. Einaudi
; REGISTRATION NUMBER: [illegible]
; REFERENCE/DOCKET NUMBER: [illegible]
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408 4000
; TELEFAX: 202 408 4400
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Escherichia [illegible]
US-08-233-617-4

Query Match [illegible]; Score [illegible]; DB 5; Length [illegible];
Best Local Similarity [illegible]; Prod. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
[illegible]

Publication Title:
Typical (*) Time Time Application #:

My Run Search, Total Searches

Run on: February 5, 2003, 12:41:51 ; Search time 4.56 Seconds
(without alignments)
24.34 Million cell updates/sec

Title: US-09-423-100-5
Perfect score: 294
Sequence: 1 EVN_HLYSHEVEAIYLV(.....)VEQ 111

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 120655 seqs, 24042267 residues

Total number of hits satisfying chosen parameters: 241310

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PCT.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PCT.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PCT.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PCT.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PCT.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US04_NEW_PCT.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US04_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US03_NEW_PCT.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US03_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US02_NEW_PCT.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US02_PUBCOMB.pep:*

Total No. is the number of results predicted by chance to have a score greater than or equal to the result being printed.
Match is Percent Similarity, and is the basis for ordering the listing.

SUMMARIES

result							
No.	Score	Match	Length	DB	ID	Description	
1	294	100.0	111	2	US-09-423-100-5	Sequence 5, Appl	
2	294	100.0	111	2	US-09-423-100-7	Sequence 7, Appl	
3	294	100.0	111	2	US-09-423-100-9	Sequence 9, Appl	
4	279.5	95.1	111	10	US-09-990-100-1	Sequence 1, Appl	
5	275.5	93.7	138	9	US-09-961-687-10	Sequence 10, Appl	
6	274	92.9	150	9	US-09-961-687-9	Sequence 9, Appli	
7	267	90.8	99	2	US-09-990-100-1	Sequence 1, Appli	
8	267	90.8	99	2	US-09-423-100-11	Sequence 11, Appli	
9	267	90.8	99	2	US-09-423-100-13	Sequence 13, Appli	
10	267	90.8	99	2	US-09-423-100-15	Sequence 15, Appli	
11	267	90.8	99	2	US-09-423-100-17	Sequence 17, Appli	
12	267	90.8	111	2	US-09-423-100-4	Sequence 4, Appli	
13	267	90.8	111	2	US-09-423-100-125	Sequence 125, App	
14	267	90.8	111	10	US-09-961-220-6	Sequence 6, Appli	
15	267	90.8	117	10	US-09-961-687-6	Sequence 6, Appli	
16	267	90.8	150	10	US-09-961-687-2	Sequence 2, Appli	
17	265.5	90.3	124	2	US-09-990-100-11	Sequence 11, Appl	
18	265.5	90.3	124	2	US-09-990-100-71	Sequence 71, Appl	
19	265.5	90.3	124	2	US-09-990-100-73	Sequence 73, Appl	
20	265.5	90.3	124	2	US-09-990-100-75	Sequence 75, Appl	
21	265.5	90.3	124	2	US-09-990-100-77	Sequence 77, Appl	
22	265.5	90.3	124	2	US-09-990-100-79	Sequence 79, Appl	
23	265.5	90.3	124	2	US-09-990-100-81	Sequence 81, Appl	
24	265.5	90.3	124	2	US-09-990-100-83	Sequence 83, Appl	
25	265	90.1	111	2	US-09-990-100-3	Sequence 3, Appli	
26	265	90.1	111	2	US-09-990-100-5	Sequence 5, Appli	
27	265	90.1	111	10	US-09-961-220-5	Sequence 5, Appli	
28	265	90.1	111	10	US-09-961-687-5	Sequence 5, Appli	
29	263.5	89.6	111	2	US-09-990-100-9	Sequence 9, Appli	
30	263.5	89.6	111	2	US-09-990-100-13	Sequence 13, Appl	
31	263.5	89.6	111	2	US-09-990-100-15	Sequence 15, Appl	
32	263.5	89.6	111	2	US-09-990-100-17	Sequence 17, Appl	
33	263.5	89.6	111	2	US-09-990-100-19	Sequence 19, Appl	
34	263.5	89.6	111	2	US-09-990-100-21	Sequence 21, Appl	
35	263.5	89.6	111	2	US-09-990-100-23	Sequence 23, Appl	
36	263.5	89.6	111	2	US-09-990-100-25	Sequence 25, Appl	
37	263.5	89.6	111	2	US-09-990-100-27	Sequence 27, Appl	
38	263.5	89.6	111	2	US-09-990-100-29	Sequence 29, Appl	
39	263.5	89.6	111	2	US-09-990-100-31	Sequence 31, Appl	
40	263.5	89.6	111	2	US-09-990-100-33	Sequence 33, Appl	
41	263.5	89.6	111	2	US-09-990-100-35	Sequence 35, Appl	
42	263.5	89.6	111	2	US-09-990-100-37	Sequence 37, Appl	
43	263.5	89.6	111	2	US-09-990-100-39	Sequence 39, Appl	
44	263.5	89.6	111	2	US-09-990-100-41	Sequence 41, Appl	
45	263.5	89.6	111	2	US-09-990-100-43	Sequence 43, Appl	

ALIGNMENTS

RESULT 1
US-09-423-100-5
; Sequence 5, Application US/09423100
; Patent No. US20020137141A1
; GENERAL INFORMATION:
; APPLICANT: Tan, Xiaojia
; TITLE OF INVENTION: Secreted Proteins, Polynucleotides
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Genomics, Inc.
; STREET: 3160 Porter Drive, Palo Alto
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,100
; FILING DATE: 03-NOV-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/107,106
; FILING DATE: 04-NOV-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PI-0192 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-853-1038
; TELEX: 910-371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 111 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENT DESCRIPTION: SEQ ID NO: 5:

RESULT 2
US/10-423-100-5
; Sequence 5, Application US/10423100
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ke
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/354,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/042,100
; FILING DATE: 13-Mar-1998
; APPLICATION NUMBER: US 08/822,100
; FILING DATE: 11-Mar-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020267-000438
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US 10/423,100-5

Query Match 100.0%; Score 294; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.9e-28;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80

RESULT 3
US/10-423-100-6
; Sequence 6, Application US/10423100
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ke
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/354,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/042,100
; FILING DATE: 13-Mar-1998
; APPLICATION NUMBER: US 08/822,100
; FILING DATE: 11-Mar-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020267-000438
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 107 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US 10/423,100-6

Query Match 100.0%; Score 294; DB 5; Length 107;
Best Local Similarity 100.0%; Pred. No. 2.9e-28;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80

RESULT 4
US/10-423-100-7
; Sequence 7, Application US/10423100
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ke
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/354,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: Unknown
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/042,100
; FILING DATE: 13-Mar-1998
; APPLICATION NUMBER: US 08/822,100
; FILING DATE: 11-Mar-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Mycroft, Frank J
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 020267-000438
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: Unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US 10/423,100-7

Query Match 100.0%; Score 294; DB 5; Length 150;
Best Local Similarity 100.0%; Pred. No. 4.1e-28;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 44 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 123

RESULT 5
US/10-423-100-5
; Sequence 5, Application US/10423100
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; Article Author, Coauthor, Title
; APPLICANT: Lowman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712P1
; CURRENT APPLICATION NUMBER: US/10/080,410
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: US/60/177,924
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 5
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US 10/080,410-5

Query Match 91.7%; Score 270.5; DB 9; Length 51;
Best Local Similarity 98.1%; Pred. No. 9.6e-27;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 EVNQHLVGSHLVEALYLVCGERGFFYTPKTRREVEDPQVEQLELGGSPGDLQTLALEVAQQKRGIVDQCCTSICSLYQLENYCN 80

RESULT 5
US-09-423-100-5
; Sequence 5, Application US/09423100
; Patent No. US6004787
; GENERAL INFORMATION:
; APPLICANT: Kjellson, Frans Bertil
; APPLICANT: Andersson, Sven
; TITLE OF INVENTION: Method for making insulin precursors and
; TITLE OF INVENTION: insulin precursors or analogues having improved fermentation
; TITLE OF INVENTION: yield in yeast
; FILE REFERENCE: 6148.4000-US
; CURRENT APPLICATION NUMBER: US/09/423,711
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: PA 1998 0[illegible]
; PRIOR FILING DATE: 2000-04-[illegible]
; PRIOR APPLICATION NUMBER: DA [illegible]
; PRIOR FILING DATE: [illegible]-12-29
; PRIOR APPLICATION NUMBER: 60/[illegible]
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.
; SEQ ID NO 5
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-423-100-5

Query Match 94.7%; Score 470.5; DB 10; Length 124;
Best Local Similarity 94.3%; Pred. No. 5.5e-26;
Matches 50; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 EVNQHLCGSHLVEALYLVCGERGFFYTPKAGIVEQCCTSICSLYQLENYCN 52
[illegible]
Db 1 EVNQHLCGSHLVEALYLVCGERGFFYTPKAGIVEQCCTSICSLYQLENYCN 52

RESULT 6
US-09-423-100-5
; Sequence 10, Application US/09423100
; Patent No. US6004787
; GENERAL INFORMATION:
; APPLICANT: Markussen, Jan
; APPLICANT: Andersen, Ib
; APPLICANT: Antonsen, Steen
; APPLICANT: Brandt, Jakob
; APPLICANT: Kollerup, Peter
; APPLICANT: Ramsoe, Bent Erik
; APPLICANT: Kaarsholm, Niels Christian
; TITLE OF INVENTION: INSULIN DERIVATIVES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk A/S, Patent Department
; STREET: 1 Novo Alle
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10154-0401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423,711
; FILING DATE: 21-May-1999
; CLASSIFICATION: [illegible]
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: [illegible]
; SEQ ID NO 10
; LENGTH: [illegible]
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-423-100-5

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 7
US-09-423-100-5
; Sequence 11, Application US/09423100
; Patent No. US6004787
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Michelle
; APPLICANT: Ullrich, Mark
; APPLICANT: Valdes, Felix
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/09/423,711
; CURRENT FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; PRIOR APPLICATION NUMBER: [illegible]
; PRIOR FILING DATE: [illegible]
; NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 5
; LENGTH: 50
; TYPE: PRT
; ORGANISM: [illegible]
US-09-423-100-5

Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
Best Local Similarity [illegible]; Pred. No. [illegible];
Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy [illegible]
Db [illegible]

RESULT 8
US-09-423-100-5
; Sequence 12, Application US/09423100
; Patent No. US6004787
; GENERAL INFORMATION:
; APPLICANT: Eli Lilly and Co.
; APPLICANT: KITABIRA, Seiichi
; APPLICANT: KOYAN, Yoshihiko
; APPLICANT: NAKAJIMA, Kazuo
; APPLICANT: NISHI, Nobuaki
; TITLE OF INVENTION: [illegible]
; FILE REFERENCE: [illegible]
; CURRENT APPLICATION NUMBER: US/09/423,711
; CURRENT FILING DATE: [illegible]

; PRIOR APPLICATION NUMBER: JP 2000-174691
; PRIOR FILING DATE: 2000-06-12
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-100-5

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 60.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 29
||| |||||||||||||||||||||||||||||
Db 1 PVNLHLCGSHLVEALYLVCGERGFFYTPKT 29

Qy 31 KGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||
Db 61 SLQKKGIVEQCCTSICSLYQLENYCN 86

RESULT 9
US-10-028-410-2
; Sequence 2, Application US/10028410
; Patent No. US20020160955A1
; GENERAL INFORMATION:
; APPLICANT: Dubaquie, Yves
; APPLICANT: Looman, Henry
; TITLE OF INVENTION: PROTEIN VARIANTS
; FILE REFERENCE: P1712R1-1
; CURRENT APPLICATION NUMBER: US/10/028,410
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: US/60/477,924
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 6
; SEQ ID NO 2
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-028-410-2

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 60.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 29
||||||||||||||||||||||||||||||
Db 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 29

Qy 31 KGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||
Db 61 SLQKKGIVEQCCTSICSLYQLENYCN 86

RESULT 10
US-10-054-873-4
; Sequence 4, Application US/10054873
; Patent No. US20020164712A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Zhong Ru
; TITLE OF INVENTION: Chimeric Protein Containing an
; Intramolecular Chaperone-Like Sequence
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/054,873
; FILING DATE: 22-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/423,100
; FILING DATE: 11-DEC-2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Ng, Janis J.
; REGISTRATION NUMBER: 46,946
; REFERENCE/DOCKET NUMBER: 018282-000100US
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 86 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-054-873-4

Query Match 90.8%; Score 267; DB 9; Length 86;
Best Local Similarity 60.5%; Pred. No. 3.7e-25;
Matches 52; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

Qy 1 PVNQHLCGSHLVEALYLVCGERGFFYTPKT 29
||||||||||||||||||||||||||||||
Db 1 PVNLHLCGSHLVEALYLVCGERGFFYTPKT 29

Qy 31 KGIVEQCCTSICSLYQLENYCN 52
|||||||||||||||||||||
Db 61 SLQKKGIVEQCCTSICSLYQLENYCN 86

RESULT 11
US-09-947-563-4
; Sequence 4, Application US/09947563
; Patent No. US20020156234A1
; GENERAL INFORMATION:
; APPLICANT: Rubroder, Franz-Josef
; Keller, Reinhold
; TITLE OF INVENTION: Improved process for obtaining
; insulin precursors having correctly bonded cystine bridges
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/947,563
; FILING DATE: 07-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/714,206
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Leslie McDonell
; REGISTRATION NUMBER: 34,872

; TITLE OF INVENTION: DNAS ENCODING NEW FUSION PROTEINS AND PROCESSES FOR
; TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE
; TITLE OF INVENTION: DNAS
; FILE REFERENCE: 08210.026
; CURRENT APPLICATION NUMBER: US/09/280,030A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/079,673
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 63
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus is
; OTHER INFORMATION: an amino acid sequence of
; OTHER INFORMATION: MSRog MSRng1 Met-Proinsulin
RS 09/280,030-63

Query Match 90.8%; Score 267; DB 10; Length 117;
Best Local Similarity 60.5%; Pred. No. 5.5e-25;
Matches 92; Conservative 0; Mismatches 0; Indels 45; Gaps 1;

Qy 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKT- 30
|||||||||||||||||||||||||||||| |
Db 1 FVNQHLCGSHLVEALYLVCGERGFFYTPKTRREAEDLQVGQVELGGGPGAGSLQPLALEG 60

Qy 31 GSLQKRGIVEQCCTSICSLYQLENYCN 57
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SLQKRGIVEQCCTSICSLYQLENYCN 87

Search completed: February 5, 2003, 12:48:48
Job time : 4.595 secs